

In silico-Design, Synthese und Evaluation von Kinase- und Proteasom-Inhibitoren sowie Fluoreszenzsonden.



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Vom Fachbereich Chemie
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines
Doctor rerum naturalium (Dr. rer. nat.)

genehmigte
kumulative Dissertation

vorgelegt von
Christoph Scholz, M.Sc.
aus Potsdam

Referent:	Prof. Dr. Boris Schmidt
Korreferenten:	Prof. Dr. Harald Kolmar Prof. Dr. Tanja Schirmeister
Tag der Einreichung:	29. April 2015
Tag der mündlichen Prüfung:	29. Juni 2015


Darmstadt 2015

D17



A good model is worth its weight in gold.

Francis Crick



Die vorliegende Arbeit wurde unter der Leitung von Herrn Prof. Dr. Boris Schmidt in der Zeit von Oktober 2012 bis Juni 2015 am Clemens-Schöpf-Institut für Organische Chemie und Biochemie der Technischen Universität Darmstadt angefertigt.

Danksagung

Besonderer Dank gilt meinem Doktorvater BORIS SCHMIDT für die Betreuung dieser Arbeit, sein spürbares Vertrauen in mich sowie die stete Unterstützung und die wissenschaftlichen Freiheiten, die ich während meiner Dissertation genoss.

STEPHANIE MÜLLER danke ich für ihre Hilfsbereitschaft – eine derart kompetente Sekretärin zu haben, ist ein echter Glücksfall für den Arbeitskreis.

Ich danke allen ehemaligen Mitgliedern des Arbeitskreises für die schöne gemeinsame Zeit, insbesondere EVA NAUMANN, ANDREA ZALL, DANIEL KIESER, ALEXANDER BOLÄNDER, THOMAS KRAMER, FABIO LO MONTE, PARUL GOEL und STEFAN GÖRING.

Den aktiven Mitarbeitern des Arbeitskreises – THERESA NEUMANN, JOHANNES PILAKOWSKI und DENNIS BENSINGER – danke ich besonders für die wunderbaren und produktiven letzten Monate im Viererteam. Die Zeit in Minimalbesetzung war stets geprägt von außerordentlicher Hilfsbereitschaft, fachlicher Kompetenz und Freundschaft.

Mein tiefster Dank gilt CONSTANTIN VOSS, mit dem ich nicht nur das Forschungsthema „Proteasom“, sondern vor allem die schönsten Momente meiner Promotion teilen durfte. Mit Freude denke ich an unsere gemeinsame Ära im Labor, unsere Reisen nach München, Berlin und Madrid sowie unbezahlbare Diskussionen, die unsere wissenschaftliche Arbeit wie kaum etwas anderes beflügelten.

SABINE KNORR danke ich für ihre unvergleichliche Art, die schöne Zeit in Amsterdam und ihre Programmierkünste.

Meinen Bachelorstudenten – CHRIS KORBIEN, PHILIP EMMERICH und ISABEL MORITZ – danke ich für ihre Hilfe und die wichtigen Erfahrungen, die ich durch die Betreuung ihrer Abschlussarbeiten sammeln konnte.

Unseren Kooperationspartnern in München – EVA HUBER, MICHAEL GROLL, MARTIN STEIN und PHILIPP BECK – sowie in Berlin – PETER-MICHAEL KLOETZEL und ULRIKE KUCKELKORN – danke ich für ihre Gastfreundschaft und den wertvollen fachlichen Austausch.

MIREILLE KRIER und PAUL CZODROWSKI danke ich für ihr motivierendes Interesse an meiner Arbeit, so manchen Denkanstoß und faszinierende Einblicke in das professionelle *Modeling*-Umfeld der Merck KGaA.

Mein herzlichster Dank gilt meiner FAMILIE, vor allem meiner Mutter SYLVIA, meinem Stiefvater FRANZ, meinem Bruder SEBASTIAN und meiner Freundin NADIN, ohne deren Vertrauen und Unterstützung meine Promotion undenkbar gewesen wäre.

Veröffentlichungen

Teilergebnisse dieser Arbeit wurden an folgenden Stellen vorgestellt, veröffentlicht oder werden zur Veröffentlichung eingereicht:

- Christoph Scholz, Sabine Knorr, Kay Hamacher, Boris Schmidt, "DOCKTITE – A Highly Versatile Step-by-Step Workflow for Covalent Docking and Virtual Screening in the Molecular Operating Environment", *Journal of Chemical Information and Modeling* **2015**, 55, 398-406.
- Binia Drung, Christoph Scholz, Valéria A. Barbosa, Azadeh Nazari, Maria H. Sarragiotto, Boris Schmidt, "Computational & experimental evaluation of the structure/activity relationship of beta-carbolines as DYRK1A inhibitors", *Bioorganic and Medicinal Chemistry Letters* **2014**, 24, 4854-4860.
- Constantin Voss, Christoph Scholz, Sabine Knorr, Philipp Beck, Martin L. Stein, Andrea Zall, Ulrike Kuckelkorn, Peter M. Kloetzel, Michael Groll, Kay Hamacher, Boris Schmidt, "alpha-Keto Phenylamides as P1'-Extended Proteasome Inhibitors", *ChemMedChem* **2014**, 9, 2557-2564.
- Alexander Boländer, Daniel Kieser, Christoph Scholz, Roland Heyny-von Haußen, Gerhard Mall, Valérie Goetschy, Christian Czech, Boris Schmidt, "Synthesis of methoxy-X04 derivatives and their evaluation in Alzheimer's disease pathology", *Neurodegenerative Diseases* **2014**, 13, 209-213.
- Christoph Scholz, Constantin Voss, Sabine Knorr, Ulrike Kuckelkorn, Kay Hamacher, Peter M. Kloetzel, Boris Schmidt, "Paradigm Caught by a Mouse Trap: 20S Proteasome's $\beta 5$ Subunit is Not Chymotrypsin-like.", wird voraussichtlich **2015** zur Veröffentlichung eingereicht.

Posterbeiträge, Vorträge und Datenbankeinträge:

- Christoph Scholz, Sabine Knorr und Boris Schmidt, „CovDock – A highly versatile step-by-step workflow for covalent docking and virtual screening in MOE.“, User Group Meeting: Chemical Computing Group, Straßburg (F), 22.-25.04.2014 (Poster).
- Christoph Scholz, „DOCKTITE – A highly versatile step-by-step workflow for covalent docking and virtual screening in MOE.“, User Group Meeting: Chemical Computing Group, London (GB), 12.-15.05.2015 (Vortrag).
- PDB-ID: 4R02, yCP in complex with BSc4999 (**2014**, RCSB PDB).^[1]

Danksagung.....	iv
Veröffentlichungen.....	v
Inhaltsverzeichnis	vi
Abkürzungsverzeichnis	viii
1. Einleitung.....	1
1.1 Das Ubiquitin/Proteasom-System	1
1.1.1 Die Ubiquitylierung von Proteinen.....	1
1.1.2 Der proteasomale Proteinabbau.....	3
1.1.3 Das Proteasom als Zielprotein in der molekularen Onkologie	6
1.1.4 Inhibitoren des 20S-Proteasoms.....	8
1.2 <i>Molecular Modeling</i>	14
1.2.1 Pharmakophor-Modellierung	20
1.2.2 Quantitative Struktur/Aktivitäts-Beziehung (QSAR).....	24
1.2.3 Kovalentes Docking.....	25
1.3 Die ALZHEIMER-Demenz	30
1.3.1 Die Amyloid-Hypothese	31
1.3.2 Die Tau-Hypothese	32
1.3.3 Diagnose durch Fluoreszenz- <i>Imaging</i>	33
1.3.4 DYRK1A-Inhibition als potentielle Therapie der AD	36
2. Zielsetzung.....	39
3. Kumulativer Teil	42
3.1 Optimierung von -Ketophenylamiden als Proteasominhibitoren	42
3.2 Untersuchung des Einflusses eines Ligand/Met45-Kontaktes auf die enzymatische Aktivität des Proteasoms.....	75
3.3 Entwicklung von DOCKTITE, einer Methode zum automatisierten Docking kovalenter Liganden	85

3.4	Entwicklung einer optimierten Syntheseroute zu Methoxy-X04-basierten Fluoreszenzfarbstoffen zur Diagnose der Alzheimer-Demenz	109
3.5	Untersuchung der Struktur/Aktivitäts-Beziehung (SAR) von β -Carbolinen als DYRK1A-Inhibitoren und Steigerung der Selektivität gegenüber MAO-A mit Hilfe moderner <i>in silico</i> -Methoden.....	126
4.	Zusammenfassung und Ausblick	182
	Literaturverzeichnis	185
	Anhang.....	194

Abkürzungsverzeichnis

a	Aktivität
abs.	Absorption
AD	ALZHEIMER-Demenz
AMC	7-Amino-4-methylcoumarin
AMP	Adenosinmonophosphat
AP	Amyloider Plaque
APP	Amyloidvorläufer-Protein (engl.: <i>Amyloid precursor protein</i>)
Arg/R	L-Arginin
ATP	Adenosintriphosphat
A β	Amyloid- β -Peptid
BrAAP	Verzweigtkettige Aminosäuren-bevorzugend (engl.: <i>Branched chain amino acid preferring</i>)
CAS	<i>Chemical Abstracts Service</i>
ChTL	Chymotrypsin-ähnlich (engl.: <i>Chymotrypsin-like</i>)
CL	Caspase-ähnlich (engl.: <i>Caspase-like</i>)
CP	Kernpartikel (engl.: <i>Core particle</i>)
CSD	<i>Cambridge Structural Database</i>
CT	Computertomographie
Cys/C	L-Cystein
Da	Dalton [g/mol]
DNA	Desoxyribonucleinsäure
E1	Ubiquitin-aktivierendes Enzym
E2	Ubiquitin-konjugierendes Enzym
E3	Ubiquitin-Ligase
EL	Elastase-ähnlich (engl.: <i>Elastase-like</i>)
em.	Emission
FDA	<i>Food and Drug Administration</i>
G	Freie Enthalpie
Gln/Q	L-Glutamin
Glu/E	L-Glutaminsäure
Gly/G	Glycin

Gly47N	N-terminales Amin von Glycin47
H-Brücke	Wasserstoffbrücke
IC ₅₀	Inhibitionskonstante (Konzentration für 50% Inhibierung)
ICD	<i>International Statistical Classification of Diseases and Related Health Problems</i>
ID	Identifikationsnummer
IFN-γ	γ-Interferon
IUPAC	<i>International Union of Pure and Applied Chemistry</i>
K _i	Dissoziationskonstante
Leu/L	L-Leucin
LFER	Lineare freie Enthalpie-Beziehung (engl.: <i>linear free energy relationship</i>)
Lys/K	L-Lysin
M	Molar [mol/L]
max.	Maximal
MCSS	<i>Multiple Copy Simultaneous Search</i>
MD	Moleküldynamik
Met/M	L-Methionin
MHC	Haupthistokompatibilitätskomplex (engl.: <i>Major histocompatibility complex</i>)
MOE	<i>Molecular Operating Environment</i>
NFT	Neurofibrillenbündel (engl.: <i>Neurofibrillary tangle</i>)
NMR	Kernmagnetresonanz (engl.: <i>Nuclear magnetic resonance</i>)
P	Phosphat
PDB	<i>Protein Data Bank</i>
PET	Positronen-Emissions-Tomographie
pH	$-\log(a(\text{H}_3\text{O}^+))$
Ph4	Pharmakophor
Phe/F	L-Phenylalanin
PHF	Paariges helikales Filament
pIC ₅₀	$-\log(\text{IC}_{50})$
Pkt.	Punkt
PN	Periphere Neuropathie

PP _i	Pyrophosphat
QSAR	Quantitative Struktur/Aktivitäts-Beziehung (engl.: <i>Quantitative structure/activity relationship</i>)
RDST	<i>Rapid Dementia Screening Test</i>
RMSD	Wurzel der mittleren quadratischen Abweichung von Atompositionen (engl.: <i>Root-mean-square deviation of atomic positions</i>)
RNS	Reaktive Stickstoffspezies (engl.: <i>Reactive nitrogen species</i>)
ROS	Reaktive Sauerstoffspezies (engl.: <i>Reactive oxygen species</i>)
RP	Regulatorischer Partikel
S	Svedberg
SAR	Struktur/Aktivitäts-Beziehung (engl.: <i>Structure/activity relationship</i>)
Ser/S	L-Serin
SNAAP	Kleine, neutrale Aminosäuren-bevorzugend (engl.: <i>Small neutral amino acid preferring</i>)
SPECT	Einzelphotonen-Emissions-Computertomographie (engl.: <i>Single-photon emission computed tomography</i>)
suc	Succinyl
Thr/T	L-Threonin
Thr1O γ	γ -Sauerstoffatom der Threonin1-Seitenkette
Thr1N	N-terminales Amin von Threonin1
TL	Trypsin-ähnlich (engl.: <i>Trypsin-like</i>)
Ub	Ubiquitin
UPS	Ubiquitin/Proteasom-System
Z	Benzyloxycarbonyl
λ	Wellenlänge [m]

1. Einleitung

1.1 Das Ubiquitin/Proteasom-System

Das Proteom, also die Gesamtheit aller Proteine in einem Organismus zu einem bestimmten Zeitpunkt, unterliegt einem dynamischen Gleichgewicht zwischen Proteinabbau und -synthese. Der Zerlegung von Proteinen in ihre Bestandteile, die Aminosäuren, wurde bis zur Mitte des 20. Jahrhunderts nur wenig Aufmerksamkeit geschenkt. 1942 stellte R. SCHOENHEIMER seine Forschungen über „Die Dynamik der Körperbestandteile“ vor, durch die er mittels Isotopenmarkierung den ständigen Auf- und Abbau von Proteinen nachweisen konnte.^[2] M.V. SIMPSON konnte elf Jahre später zeigen, dass es neben den bekannten, ATP-unabhängigen Abbauwegen durch extrazelluläre Proteasen, wie Trypsin und Chymotrypsin, energieverbrauchende Prozesse geben muss, die für den Abbau intrazellulärer Proteine verantwortlich sind.^[3] Besonderes Interesse wurde daraufhin dem im gleichen Zeitraum von C. DE DUVE erstmals beschriebenen lysosomalen Abbau von Proteinen zuteil.^[4] Lysosomen sind Zellorganellen, die zelleigene und zellfremde Stoffe durch Endozytose aufnehmen und diese in ihrem Inneren durch eine Vielzahl von Verdauungsenzymen in ihre Bestandteile zerlegen können. Um die Zelle im Falle des Austretens dieser Enzyme zu schützen, funktionieren diese nur bei einem stark erniedrigten pH-Wert von 4,5 - 5 und sind im pH-Milieu des Zytoplasmas inaktiv. Zum Aufrechterhalten dieses Protonengradienten benötigen Lysosomen Energie, welche sie aus der Speicherform ATP beziehen.^[5] Anfang der 80er-Jahre entdeckten und charakterisierten die Mediziner A. CIECHANOVER, A. HERSHKO, und I. ROSE den durch Ubiquitinmarkierung hochselektiven und ATP-abhängigen Abbau intrazellulärer Proteine durch das Ubiquitin/Proteasom-System (UPS) und wurden für ihre herausragende Forschungsleistung im Jahr 2004 mit dem Nobelpreis für Chemie ausgezeichnet.^[6]

1.1.1 Die Ubiquitylierung von Proteinen

Ubiquitin (Ub) ist mit einer Masse von 8 kDa ein relativ kleines Protein, das aus 76 Aminosäuren aufgebaut ist. Seinen Namen verdankt es seinem „ubiquitären“ Vorkommen in nahezu allen Organismen. Ub kann über seinen C-Terminus

isozeptidisch an Lysin-Seitenketten von Substratproteinen gebunden werden und dient auf diese Weise entweder als Signal zum Abbau markierter Proteine über das UPS oder reguliert wichtige Prozesse, wie Zellzykluskontrolle, DNA-Reparatur und Apoptose. Die Ubiquitinylierung erfolgt über eine Kaskade dreier Enzymreaktionen (Abbildung 1).^[7]

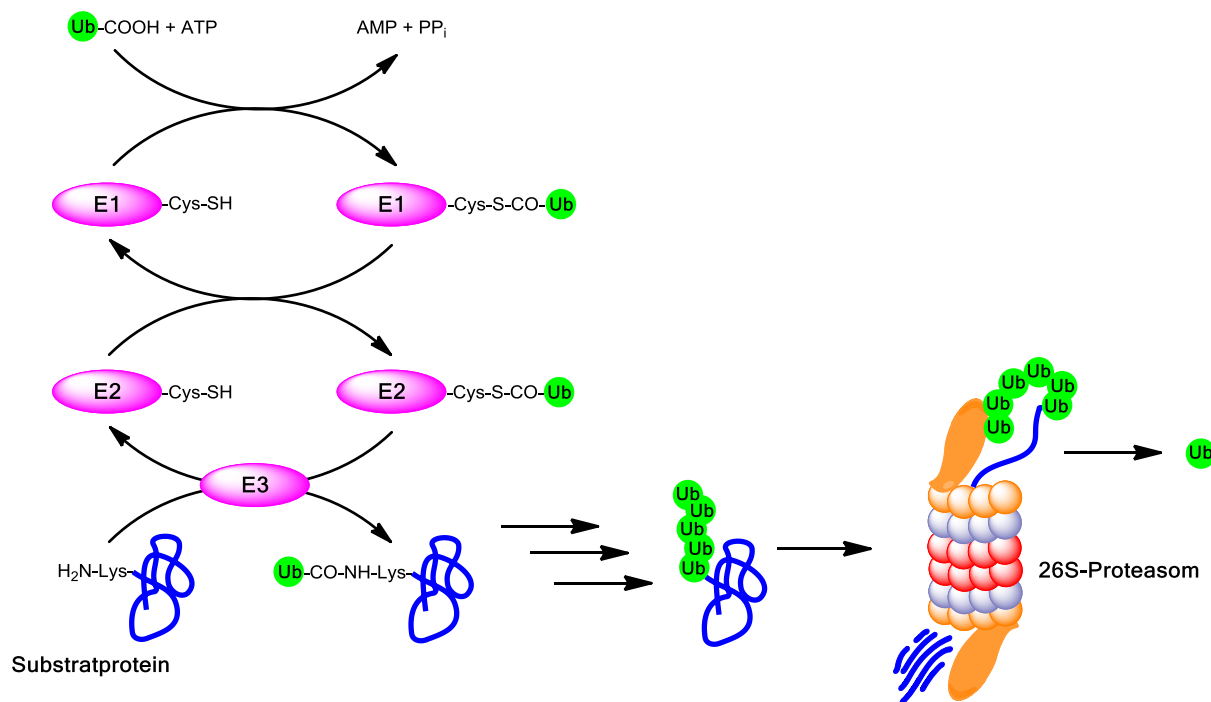


Abbildung 1: Das Ubiquitin/Proteasom-System: Schematische Darstellung der Polyubiquitinylierung eines Substratproteins über die ATP-abhängige E1/E2/E3-Enzymkaskade mit anschließender Erkennung und Abbau des Substratproteins durch das 26S-Proteasom. Modifizierte Abbildung nach BIEL et al.^[7]

Im ersten Schritt wird Ub über eine hochenergetische Thioesterbindung auf eine Cystein-Seitenkette des Ub-aktivierenden Enzyms (E1) unter ATP-Verbrauch übertragen. Im zweiten Schritt erfolgt die Umesterung auf eine Cystein-Seitenkette des Ub-konjugierenden Enzyms (E2). Durch eine für das jeweilige Substratprotein hochspezifische Ub-Ligase (E3) wird im finalen Schritt die Übertragung des Ubiquitins auf eine Lysin-Seitenkette des Zielproteins realisiert. Ub selbst verfügt über zwei exponierte Lysin-Seitenketten (Lys48 und Lys63), wodurch sich die Reaktionskaskade mehrfach wiederholen kann. Dies führt schließlich zu einer Polyubiquitinylierung des Substratproteins, welche als Signal für den proteasomalen Abbau dient.^[7]

1.1.2 Der proteasomale Proteinabbau

Das 26S-Proteasom ist ein multikatalytischer Enzymkomplex mit einer Masse von 2,5 MDa. Er besteht aus zwei regulatorischen 19S-Untereinheiten (RP) sowie dem fassartigen Kernkomplex (CP), dem 20S-Proteasom. Während die 19S-Untereinheiten für die Erkennung markierter Substratproteine, deren Entfaltung und Transport in das Innere des Substratkanals unter ATP-Verbrauch zuständig sind, beherbergt der CP die aktiven Zentren der Untereinheiten $\beta 1$, $\beta 2$ und $\beta 5$ (Abbildung 2).^[8, 9]

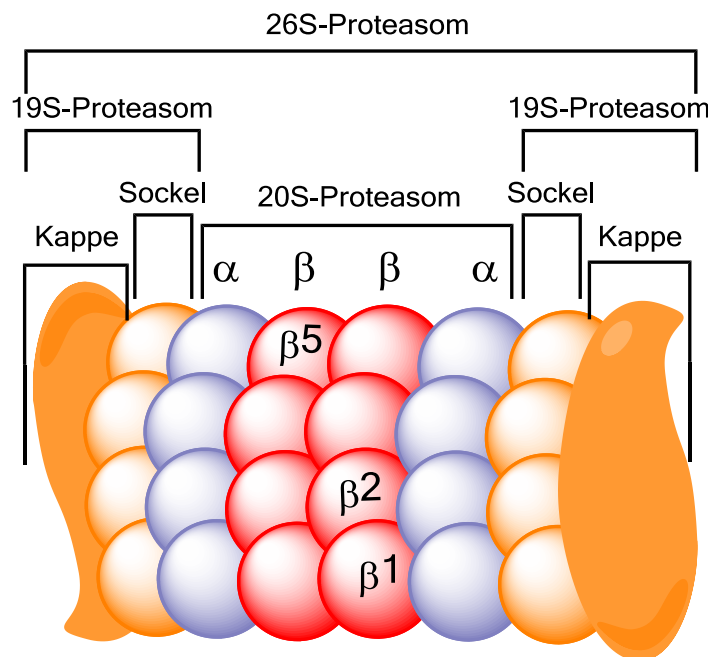


Abbildung 2: Schematische Darstellung des 26S-Proteasoms: Der Kernpartikel (CP), das 20S-Proteasom, besteht aus je zwei heptameren α - und β -Ring. Die β -Untereinheiten beherbergen die aktiven Zentren ($\beta 1$, $\beta 2$ und $\beta 5$) mit jeweils unterschiedlichen Substratspezifitäten. Der CP ist auf beiden Seiten vom sogenannten regulatorischen Partikel (RP), dem 19S-Proteasom, umgeben, welcher für die Erkennung, die Entfaltung und den Transport polyubiquitinylierter Substratproteine in das Innere des CP verantwortlich ist. Modifizierte Abbildung nach Voss.^[10]

Das 20S-Proteasom hat eine Masse von ~ 700 kDa und ist zylindrisch aufgebaut mit einer Länge von ~ 160 Å und einem Durchmesser von ~ 120 Å. Es besteht aus je zwei heptameren α - und β -Ring, woraus sich eine 72-Punkt-Symmetrie ergibt.^[9] Der CP beherbergt insgesamt sechs aktive Zentren mit drei unterschiedlichen Substratspezifitäten und einem für Proteasen seltenen Hydrolysemechanismus: Ein N-terminales Threonin (Thr1) bildet die nukleophile Seitenkette, dessen

Hydroxygruppe die Amidbindung nukleophil angreift, während das terminale Amin ein Wassermolekül sowohl koordiniert als auch aktiviert (Abbildung 3).^[11]

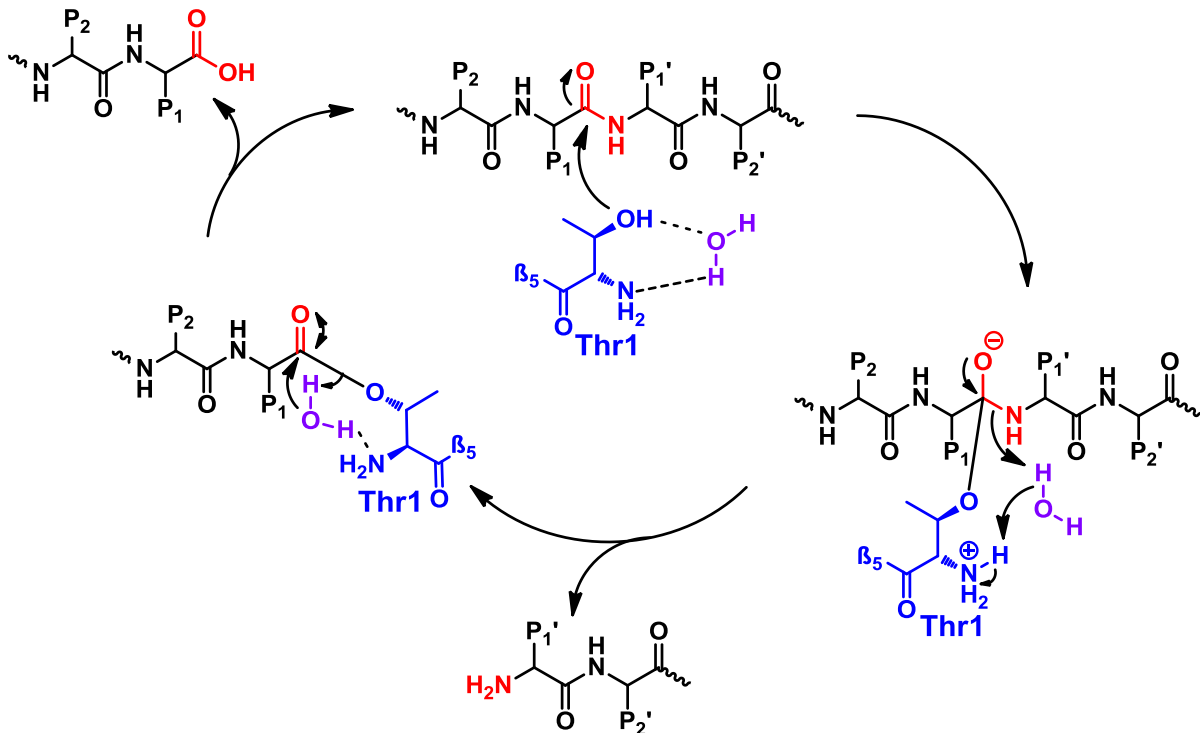


Abbildung 3: Hydrolysemechanismus eines Substratproteins durch das Proteasom: Thr1 (blau) greift über die aktivierte Hydroxygruppe die Peptidbindung eines Substratproteins (rot) nukleophil an. Nach Freisetzung des N-terminalen Peptidfragments aus dem tetraedrischen Intermediat und anschließender Freisetzung der carboxyterminalen Komponente durch Hydrolyse des Acylesters wird das katalytisch aktive Thr1 regeneriert. Modifizierte Abbildung nach KISSELEV et al.^[11]

Im ersten Schritt kommt es zu einem nukleophilen Angriff der Hydroxylgruppe des N-terminalen Thr1 (Thr1Oγ) auf die zu schneidende Peptidbindung des Substratproteins. Der Amin-Terminus (Thr1N) ermöglicht diesen Angriff, indem durch Koordinierung eines Wassermoleküls die Nukleophilie der Hydroxylgruppe signifikant erhöht wird. Im folgenden Schritt wird das entstandene Acylester-Intermediat hydrolysiert und Thr1 für den nächsten Katalysezyklus regeneriert.

Die Substratspezifität kommt durch die Struktur der Substratbindetaschen des aktiven Zentrums zustande, welche, von der Schnittstelle ausgehend, in N-terminaler Richtung des Substrates mit S1-S4 und in C-terminaler Richtung mit S1'-S4' bezeichnet werden (Abbildung 4).^[12] Nach heutigem Kenntnisstand sind primär die rot markierten Substratbindetaschen S1 und S3 für die Selektivität entscheidend,

während die S2-Bindetasche groß genug ist, um selbst sterisch anspruchsvollste Seitenketten des Substrates aufzunehmen.^[9]

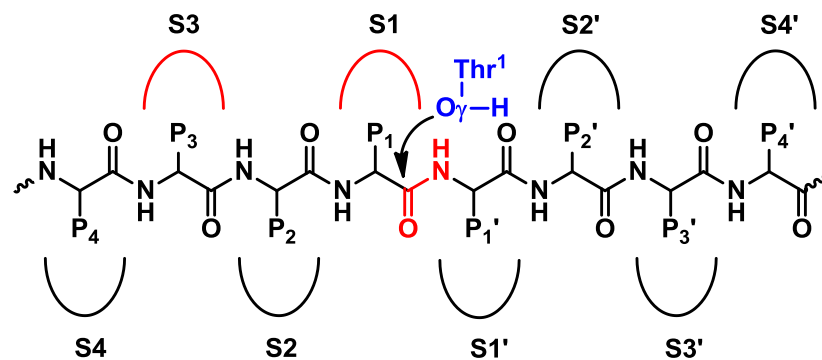


Abbildung 4: Nomenklatur des Substratbindekanals: Man unterscheidet zwischen einer gestrichenen Seite, deren Bindetaschen in C-terminaler Richtung, ausgehend von der zu schneidenden Peptidbindung, mit S1'-S4' bezeichnet werden und einer nicht-gestrichenen Seite, deren Bindetaschen in N-terminaler Richtung mit S1-S4 benannt werden.

Die Substratspezifität der aktiven Untereinheiten lässt sich verlässlich aus der Struktur der S1-Bindetasche vorhersagen, welche vor allem durch die Aminosäure an Position 45 definiert wird. Arg45 der β 1-Untereinheit ist besonders gut geeignet, Wechselwirkungen mit sauren Aminosäuren einzugehen und führt somit zu der beobachteten Caspase-ähnlichen Aktivität (CL). Die S1-Bindungstasche der β 2-Untereinheit ist durch Gly45 relativ groß. Durch Wechselwirkungen des Substrates mit der nahegelegenen Seitenkette von Glu53 werden hier große und basische P1-Reste bevorzugt, was zur Klassifizierung als Trypsin-ähnliche Aktivität (TL) führt. Die Chymotrypsin-ähnliche Aktivität (ChTL) der β 5-Untereinheit kommt durch die konformationell überaus bewegliche Seitenkette von Met45 zustande. Diese wichtige katalytische Untereinheit hydrolysiert dementsprechend nach neutralen, hydrophoben Aminosäuren. Über diese drei klassischen Substratpräferenzen hinaus wurden proteasomale Aktivitäten gegenüber verzweigtkettigen Aminosäuren (BrAAP) und kleinen neutralen Aminosäuren (SNAAP) nachgewiesen, welche nach heutigen Erkenntnissen ebenfalls der β 5-Untereinheit zugeschrieben werden können.^[9, 13, 14]

Das Proteasom arbeitet prozessiv. Ein Substratprotein verlässt den Enzymkomplex demnach nicht, bis es vollständig zu Fragmenten mit einer Länge zwischen drei und 25 Aminosäuren abgebaut wurde.^[15] Die durchschnittliche Länge der Spaltprodukte beträgt sieben bis acht Aminosäuren und ist unabhängig von der Anzahl umgebender

aktiver Zentren im Proteasom, was durch Mutationsstudien belegt werden konnte.^[15] Die hierbei erzeugten Oligopeptide werden zum Teil durch Peptidasen hydrolysiert und dem Organismus als Aminosäuren wieder zur Verfügung gestellt. Sie können jedoch auch als T-Zell-Antigene in MHC-Klasse I-Komplexen dienen. Optimale Substrate für MHC I-Komplexe sind acht bis neun Aminosäuren lang und besitzen einen basischen oder hydrophoben C-Terminus, entsprechen also Spaltprodukten, die durch die β 2- oder β 5-Untereinheit generiert werden. Um im Infektionsfall die Menge an prozessierten MHC I-Antigenen zu erhöhen, besitzen Säugetiere eine weitere Form des Proteasoms: das Immunoproteasom. Detektieren Immunzellen Infektionen, so schütten sie zur Intensivierung der Immunantwort unter anderem das Zytokin γ -Interferon (IFN- γ) aus. IFN- γ wiederum führt zu einer vermehrten Expression des Immunoproteasoms, welches sich in der Struktur der aktiven Zentren vom konstitutiven Proteasom unterscheidet. Die β 1-Untereinheit erfährt hierbei einen Aminosäureaustausch, der ihre Substratspezifität nahezu umkehrt: Durch einen Arg45Leu-Austausch von CL zu ChTL. Ein weiterer Unterschied, der durch die von M. GROLL et al. veröffentlichte Kristallstruktur des Immunoproteasoms erkannt werden konnte, ist die Größe der S1-Bindetasche der β 5-Untereinheit.^[16] Diese Veränderung kommt nicht durch einen Austausch von Met45 zustande, sondern durch einen Ser53Gln-Austausch. Die sich im Immunoproteasom befindliche Seitenkette Gln53 bewirkt eine Konformationsänderung von Met45 und vergrößert somit die S1-Bindungstasche signifikant.

1.1.3 Das Proteasom als Zielprotein in der molekularen Onkologie

Das UPS reguliert eine Vielzahl physiologischer Prozesse und wird häufig als „Recyclinghof der Zelle“ bezeichnet. Bis zu 30% der neu synthetisierten Proteine werden durch diese Qualitätskontrolle einer gesunden Zelle als schadhaft identifiziert, ubiquitinyliert und vom Proteasom abgebaut.^[17] Tumorzellen sind durch ihre hohe Teilungsrate und ihren erhöhten Metabolismus besonders anfällig für eine verminderte Proteasomaktivität. Sie „ersticken“ sozusagen an ihren eigenen Abfallprodukten. Der Einfluss einer Proteasominhibition auf das Tumorwachstum lässt sich allerdings auch detaillierter beschreiben (Abbildung 5).^[7]

spezialisierte Ub-Ligase (E3), führt zur Ubiquitinylierung von p53 und leitet somit dessen Abbau durch das Proteasom ein. In gesunden Zellen ist diese Labilität von p53 vorteilhaft, um nach erfolgter DNA-Reparatur den Zellzyklus wieder zu starten; in Tumorzellen führt der proteasomale Abbau jedoch zu einer Verminderung der p53-Akkumulation und somit zu unkontrolliertem Wachstum. NF-κB spielt als Transkriptionsfaktor ebenfalls eine entscheidende Rolle in der Krebsentstehung. Unter zellulärem Stress phosphoryliert die IκB-Kinase ihr Substrat, den NF-κB-Inhibitor IκB, und leitet dessen Ubiquitinylierung durch eine entsprechende E3-Ligase ein.^[7] Durch den Abbau von IκB wird NF-κB aktiviert, wandert in den Zellkern und erhöht dort als Transkriptionsfaktor die Expression bestimmter Gene, deren Produkte für die Regulierung verschiedenster Signalwege verantwortlich sind. Im Endeffekt kann dies zu einem Überleben der Zelle trotz eines DNA-Schadens führen. Besonders der Einfluss auf NF-κB-vermittelte Signalwege macht die Inhibierung des Proteasoms nicht nur für die Tumorthapie, sondern auch für zahlreiche Autoimmunerkrankungen zu einer interessanten Therapiestrategie.^[9]

1.1.4 Inhibitoren des 20S-Proteasoms

Als Anfang der 90er-Jahre das Interesse an der Erforschung des Proteasoms stieg, war zunächst nicht bekannt, dass es sich um eine völlig neue Klasse von Proteasen handelte. Die ersten Experimente wurden daher mit einer Vielzahl bekannter Inhibitoren für Serin- und Cysteinproteasen durchgeführt. Auf diese Weise gelang die Identifizierung **peptidischer Aldehyde**, die auch gegenüber dem CP eine inhibitorische Aktivität aufweisen. Erst die im Jahr 1995 durch W. BAUMEISTER veröffentlichte Kristallstruktur des Calpain-Inhibitors I 1 im Komplex mit dem CP des Archaeon *Thermoplasma acidophilum* bewies, dass es sich um eine Threonin-Protease handelt, deren aktives Threonin mit dem Inhibitor kovalent und reversibel ein Halbacetal ausbildet (Abbildung 6A).^[19, 20] Die Aufklärung der Quartärstruktur der aktiven Zentren ermöglichte erstmals ein rationales Inhibitordesign, um Substanzen mit selektiver Affinität für das CP zu entwickeln und mit ihnen gezielt bestimmte aktive Untereinheiten zu inhibieren. Die Substitution der Norleucin- durch eine Leucin-Seitenkette an der P1-Position und der N-terminalen Acetyl- durch die Z-Schutzgruppe, führte zu dem noch heute vielfach *in vitro* angewendeten MG-132 2.

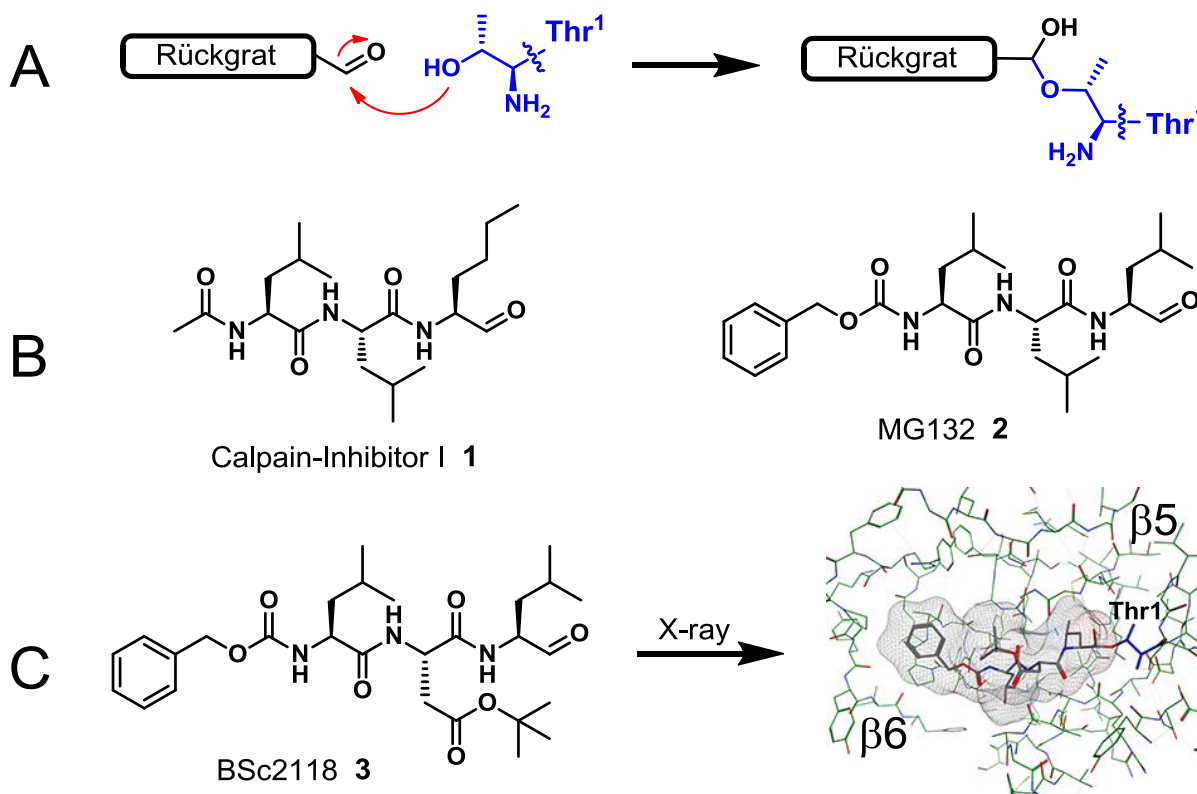


Abbildung 6: (A) Allgemeiner Bindungsmodus peptidischer Aldehyde. Durch Addition des Thr1O γ an die Carbonylgruppe des Aldehyden bildet sich ein Halbacetal aus. (B) Calpain-Inhibitor I 1 und dessen Weiterentwicklung MG132 2. (C) Der hochaktive und -selektive aldehydische Inhibitor BSc2118 3 und dessen durch röntgenkristallographische Untersuchungen aufgeklärter Bindungsmodus in der $\beta 5$ -Untereinheit des Proteasoms. (Bisher unveröffentlichte Kristallstruktur, aufgenommen und freundlicherweise zur Verfügung gestellt von M. GROLL, visualisiert mit MOE2014.09).

Der im Arbeitskreis B. SCHMIDT entwickelte tripeptidische Aldehyd BSc2118 3 ist bis heute einer der aktivsten und selektivsten Vertreter dieser Substanzklasse.^[21] Das Besondere an BSc2118 3 ist das sterisch anspruchsvolle *tert*-Butyl-aspartat an Position P2. Wie bereits in Abschnitt 1.1.2 erwähnt, weist das 20S-Proteasom keine spezifischen S2-Bindungstaschen auf, da der P2-Rest des Liganden in den inneren Kanal des Kernkomplexes hineinragen kann (Abbildung 6C). Dieser strukturelle Unterschied zu anderen Proteasen verleiht Proteasominhibitoren mit raumfüllenden P2-Resten eine hohe Selektivität.

Die erste Weiterentwicklung der Peptidaldehyde waren **peptidische Vinylsulfone**, wie das für mechanistische Studien eingesetzte, isotope markierte, NL₃VS 4 (Abbildung 7). Diese kovalent und irreversibel an das Proteasom bindende Substanzklasse ist als Michael-Akzeptor besonders affin für weiche Nukleophile wie

beispielsweise Thiole. Während Serinproteasen von Vinylsulfonen kaum inhibiert werden, zeigen sie eine hohe Reaktivität gegenüber den meisten Cysteinproteasen, wie Cathepsinen, was sich für eine Anwendung *in vivo* nachteilig auswirkt.^[9]

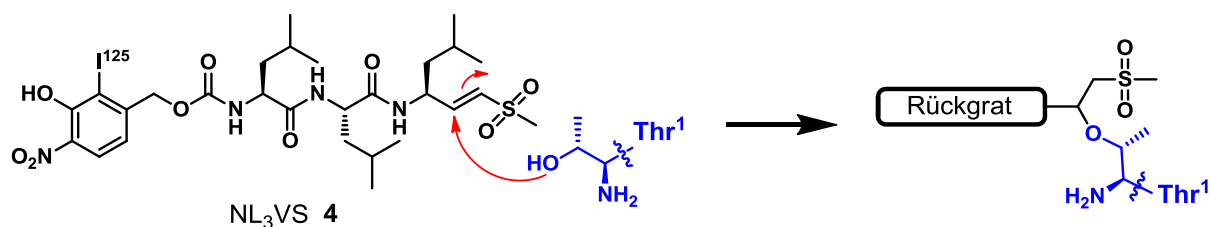


Abbildung 7: Peptidvinylsulfone, wie das mit Iod125 radioaktiv markierte NL₃VS **4**, binden mit ihrer Michael-Akzeptor-Position irreversibel an Thr1O γ und eignen sich somit insbesondere für mechanistische Studien.

Die **Boronsäure-Kopfgruppe** zeigt ein, im Vergleich zu Vinylsulfonen, inverses Selektivitätsmuster. Sie ist hochreaktiv gegenüber Sauerstoff-Nukleophilen (wie Serin oder Threonin), wird jedoch aufgrund der schwachen Bor/Schwefel-Bindung von Cysteinproteasen deutlich seltener angegriffen. Die im Vergleich zu aldehydischen Inhibitoren erhöhte Reaktivität gegenüber dem Proteasom wird durch den Vergleich der Inhibitionskonstanten (K_i) von Bortezomib **5** ($K_i = 0,62$ nM) und dessen Aldehydanalogon **6** ($K_i = 1.600$ nM) deutlich.^[22] Thr1O γ bildet mit Boronsäuren einen Boronsäureester aus – ein aus chemischer Sicht reversibler Prozess (Abbildung 8). Die Dissoziationsraten sind teilweise jedoch so gering, dass ein pseudoirreversibles Verhalten von Bortezomib **5** beobachtet wird.^[23] Aufgrund der hohen inhibitorischen Aktivität gegenüber dem CP, der Selektivität und metabolischen Stabilität schaffte es Bortezomib **5** bis in die klinischen Studien und wurde 2008 von der FDA als erster Proteasominhibitor zur Behandlung zweier äußerst schwerer Tumorerkrankungen zugelassen: des Multiplen Myeloms und des Mantelzell-lymphoms. Obwohl es in bis zu 52% der Fälle zu einer schweren Nebenwirkung – der peripheren Neuropathie (PN) – kommt, wird es seitdem weltweit zur Behandlung dieser tödlich verlaufenden Tumorerkrankungen eingesetzt.^[24, 25] PN resultiert aus einer Schädigung des peripheren Nervensystems, die mit Taubheitsgefühl in Armen und Beinen einhergeht und in schweren Fällen ($\sim 14\%$) zu vollständiger Immobilität des Patienten führt.

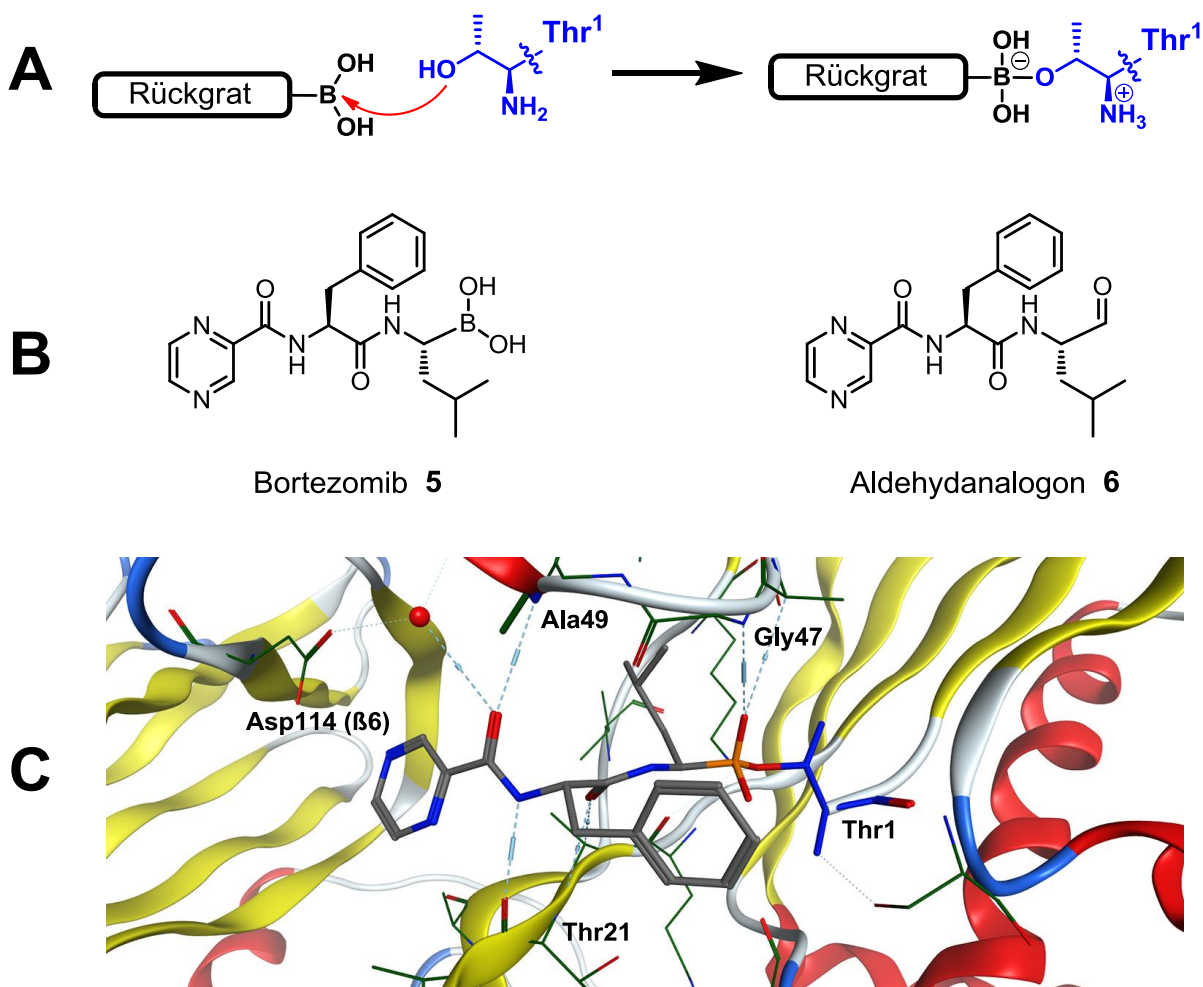


Abbildung 8: (A) Allgemeiner Bindungsmodus peptidischer Boronsäuren. Durch Addition des Thr1Oγ an das elektronendefizitäre Boratom bildet sich ein Boronsäureester-Addukt. (B) Der erste FDA-zugelassene Proteasominhibitor Bortezomib **5** und sein Aldehydanalogon **6**. (C) Bindungsmodus von Bortezomib **5** in der β5-Untereinheit des Proteasoms (PDB-ID: 2f16, visualisiert mit MOE2014.09).^[26]

Eine mögliche Ursache der PN ist die Ko-Inhibition der mitochondrialen Serinprotease HTRA2/Omi – ein sogenannter *off-target*-Effekt.^[27] HTRA2/Omi wird bei Zellstress vermehrt exprimiert und schützt vor Neurodegeneration.^[28] Dank der kombinatorischen Behandlung mit Bortezomib **5**, Chemotherapeutika, Immunmodulatoren und einer Stammzellentransplantation, konnte die Lebenserwartung der Patienten auf acht Jahre verlängert und damit seit 1990 verdoppelt werden.^[29]

Verbindungen mit **α,β-Epoxyketon-Kopfgruppe** gehören zur Gruppe der bivalenten Inhibitoren. Sie gehen sowohl mit Thr1Oγ als auch mit dem terminalen Thr1N eine kovalente Bindung ein und können auf diese Weise thermodynamisch günstige Ringsysteme bilden (Abbildung 9). Sie zeigen neben hoher Selektivität für das

20S-Proteasom außergewöhnlich niedrige Inhibitionskonstanten. Durch Optimierung des peptidischen Rückgrates des Naturstoffs Epoxomicin **7** wurde Carfilzomib **8** unter dem Arbeitsnamen PR-171 von Onyx Pharmaceuticals entwickelt.

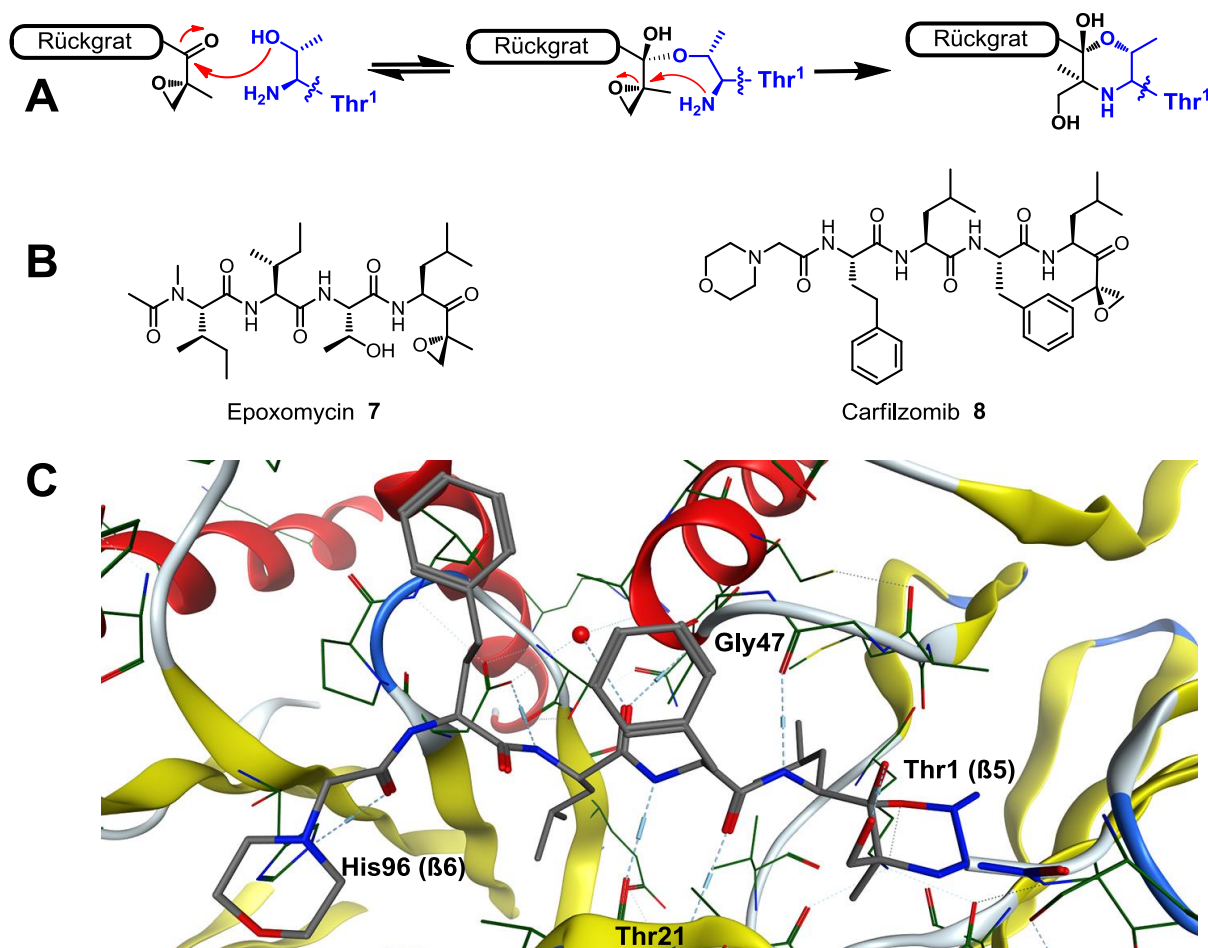


Abbildung 9: (A) Allgemeiner Bindungsmodus peptidischer α,β -Epoxyketone. Auftaktschritt ist eine reversible Addition des Thr1OH an die Carbonylgruppe, unter Bildung eines Halbacetals. Die anschließende Öffnung des Epoxids verläuft irreversibel unter Bildung eines thermodynamisch stabilen Morpholin-Rings. (B) Der Naturstoff Epoxomicin **7** und der FDA-zugelassene Epoxyketon-Inhibitor Carfilzomib **8**. (C) Bindungsmodus von Carfilzomib **8** in der $\beta 5$ -Untereinheit des Proteasoms (PDB-ID: 4qw4, visualisiert mit MOE2014.09).^[30]

In klinischen Studien wurde die Wirksamkeit von Carfilzomib **8** bei Bortezomib 5-resistentem oder wiederkehrendem Multiplen Myelom nachgewiesen, wobei periphere Neuropathien generell weniger schwer (Grad 1-2) verliefen und ihr Auftreten auf 12% der Fälle gesenkt werden konnte.^[31-33] Eine Ko-Inhibition von HTRA2/Omi wird durch Carfilzomib **8** nicht beobachtet.^[27] Carfilzomib **8** wurde im Juli 2012 von der FDA zur Behandlung rezidiver oder resistenter Formen des Multiplen Myeloms zugelassen.

Eine Inhibitor­klasse mit einem einzigartigen Bindungsverhalten stellen α -Ketophenylamide dar – wie das im Arbeitskreis B. SCHMIDT entwickelte Derivat BSc2189 **9** (Abbildung 10).^[21]

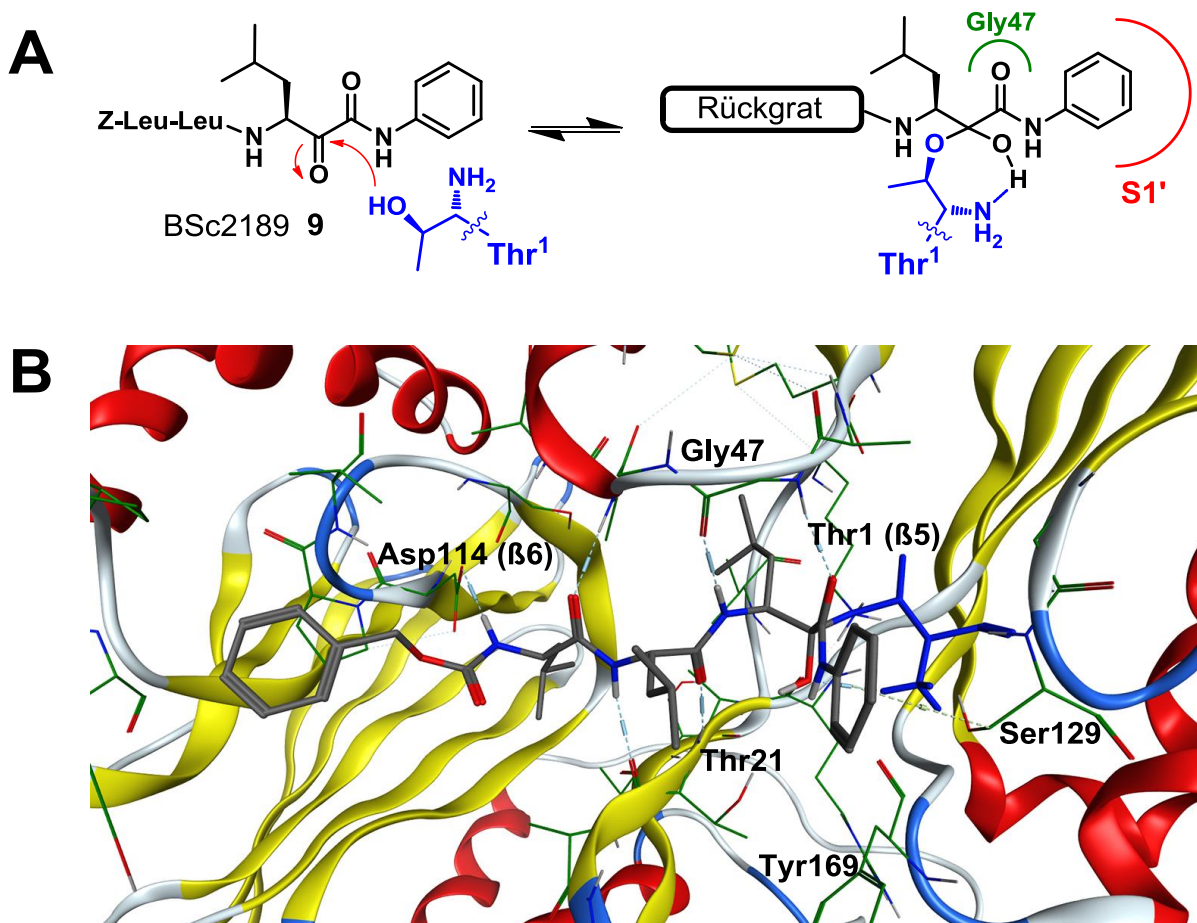


Abbildung 10: (A) Bindungsmodus des α -Ketophenylamids BSc2189 **9**, mit S1'-Kontakt (rot). Thr1O γ addiert an die Ketogruppe des Inhibitors unter Bildung eines Halbacetals, während die naszierende Hydroxygruppe vom Thr1N stabilisiert wird. Normalerweise wird diese Aufgabe durch die sogenannte „Oxyanion-Tasche“ übernommen, die im Falle der α -Ketophenylamide mit der Amid-Carbonylgruppe wechselwirkt (grün). (B) Bindungsmodus von BSc2189 **9** in der β 5-Untereinheit des Proteasoms (PDB-ID: 4no8, visualisiert mit MOE2014.09).^[23]

Das Besondere an dieser, erstmalig 1999 von S. CHATTERJEE als aktiv gegenüber dem 20S-Proteasom identifizierten, Substanzklasse ist ein Kontakt mit der gestrichenen Seite des Substratbindungskanal.^[34] Außerdem ist eine intrinsische Stabilisierung der bei der Halbacetalbildung naszierenden Hydroxygruppe durch Thr1N, statt durch die „Oxyanion-Tasche“ des Rezeptors (Gly47) auffällig. Diese wechselwirkt stattdessen mit der Carbonylgruppe des Amidterminus über eine H-Brücke, ausgehend von Gly47N. In einem systematischen Vergleich elektrophiler *Warheads* bezüglich

Aktivität, Zellgängigkeit, Reversibilität und Selektivität wurde die α -Ketophenylamid-Kopfgruppe als besonders vielversprechend für das zukünftige Design von Proteasom-Inhibitoren identifiziert.^[23] Hierzu wurden Wirkstoffe mit identischem peptidischem Rückgrat (Z-Leu₃) und sechs verschiedenen elektrophilen Kopfgruppen (Aldehyd, Boronsäure, α -Ketoaldehyd, Vinylsulfon, α,β -Epoxyketon und α -Ketophenylamid) unter identischen Assay-Bedingungen untersucht, um einen wissenschaftlichen Vergleich zu ermöglichen. Obwohl das entsprechende α,β -Epoxyketon ($IC_{50,\beta5} = 15$ nM) und die Boronsäure ($IC_{50,\beta5} = 27$ nM) niedrigere IC_{50} -Werte als BSc2189 **9** ($IC_{50,\beta5} = 70$ nM) zeigen, verhindert ihr (pseudo)-irreversibler Bindungsmodus den Einsatz als Wirkstoff gegen solide Tumoren.^[35] Die besondere Aktivität und Selektivität von BSc2189 **9** wird dem Kontakt mit der gestrichenen Seite des Substratbindungskanal zugeschrieben, weshalb es sich hervorragend als Leitstruktur für die Entwicklung neuer, S1'-selektiver, CP-Inhibitoren eignet.

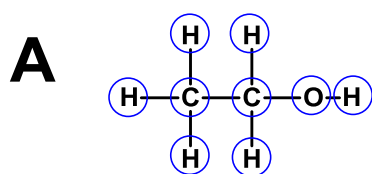
1.2 *Molecular Modeling*

1875 publizierte der damals 23-jährige J. H. VAN'T HOFF seine bahnbrechenden Theorien zur Asymmetrie des Kohlenstoffatoms und setzte somit den Grundstein für die moderne Stereochemie.^[36] Seine theoretischen Erklärungen veranschaulichte er mit Molekülmodellen, deren Herstellung aus Pappkarton er in einem siebenseitigen Appendix detailliert beschrieb. Nicht zuletzt waren es diese Modelle, die J. WISLICENUS – damals Professor für Chemie an der Universität Würzburg – zu einem enthusiastischen Verfechter der VAN'T HOFF'schen Ausführungen machten. Er orderte nicht nur größere Mengen der Modelle für seine Studenten und ließ seinen Assistenten F. HERRMANN eine deutsche Übersetzung anfertigen, sondern verteidigte die Arbeiten entschieden gegenüber der heftigen und haltlosen Kritik des angesehenen deutschen Chemikers H. KOLBE.^[37, 38] Seit diesen frühen Anfängen ist die Erstellung akkurater dreidimensionaler Modelle für die Vorhersage physikalischer, chemischer und biologischer Eigenschaften von Molekülen von enormer Bedeutung und Grundvoraussetzung für ein rationales Molekül- und Wirkstoffdesign.

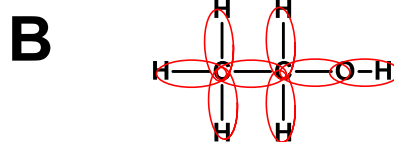
Kamen in den 1950er- und 1960er-Jahren noch hauptsächlich Modelle aus Kunststoff oder Metall zum Einsatz – J. D. WATSON und F. H. C. CRICK bauten ihre berühmte DNA-Doppelhelix in mühevoller Handarbeit aus Aluminium –, so wird modernes

Molecular Modeling vornehmlich computerbasiert durchgeführt.^[39] Heute ist darunter die Berechnung, Darstellung und Bearbeitung der dreidimensionalen Struktur von Molekülen und ihrer physikochemischen Eigenschaften zu verstehen.^[40] Im Folgenden wird eine Auswahl der wichtigsten Methoden vorgestellt.

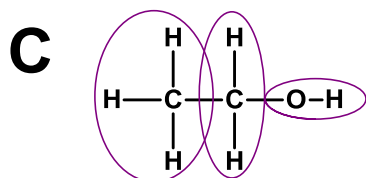
Wissensbasierte Methoden: In den letzten Jahrzehnten hat sich eine schier endlose Menge an experimentellen Daten zu Molekülstrukturen, physikalischen Eigenschaften, Proteinstrukturen, Proteinsequenzen, Struktur/Wirkungs-Beziehungen und vielem mehr angesammelt. Dieses Wissen effizient für die Beantwortung neuer Fragestellungen zu nutzen ist eine der gebräuchlichsten Methoden des *Molecular Modeling*.^[40] Zur Modellierung physikochemischer Eigenschaften haben sich hierbei besonders die Inkrementmethoden etabliert, die auf dem sog. Additionsprinzip beruhen.^[41] Sie basieren auf der Annahme, dass sich die Eigenschaft X eines Moleküls aus Beiträgen seiner Fragmente $X_{(\text{Fragment})}$ zu dieser Eigenschaft zusammensetzt. Fragmente können hierbei sowohl Atom- als auch Bindungsfragmente, funktionelle Gruppen oder größere Strukturfragmente sein (Abbildung 11). Generell gilt: Je komplexer die Fragmente, deren Eigenschaft experimentell bestimmt wurde, desto genauer kann die Eigenschaft des Zielmoleküls im Computer modelliert werden. Natürlich nimmt dadurch auch der experimentelle Aufwand deutlich zu. Nutzt man lediglich Atomfragmente der am häufigsten vorkommenden Atome Kohlenstoff, Wasserstoff, Stickstoff, Sauerstoff, Fluor, Chlor, Brom, Iod, Schwefel und Phosphor, so lassen sich bereits Hunderttausende der derzeit etwa 94 Millionen CAS-registrierten Verbindungen modellieren – und das mit nur zehn Fragmenten, deren Beitrag zu der gewünschten Eigenschaft experimentell bestimmt werden muss.^[42] Die Anwendungsgebiete der Inkrementmethoden zur Modellierung verschiedenster Eigenschaften, sog. Deskriptoren, sind vielfältig. Sie reichen von der Berechnung der Löslichkeit^[43] über die Berechnung der Verdampfungsenthalpie und -entropie^[44] bis hin zur Vorhersage der chemischen Verschiebung in NMR-Spektren.^[45]



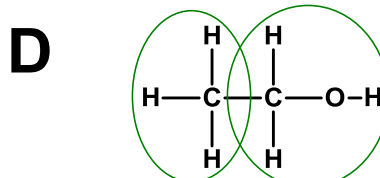
$$X = 2 \cdot X_{(C)} + 6 \cdot X_{(H)} + X_{(O)}$$



$$X = X_{(C-C)} + 5 \cdot X_{(C-H)} + X_{(C-O)} + X_{(O-H)}$$



$$X = X_{(CH_3)} + X_{(CH_2)} + X_{(OH)}$$



$$X = X_{(CH_3)} + X_{(CH_2OH)}$$

Abbildung 11: Schematische Zerlegung eines Ethanol-Moleküls in **(A)** Atomfragmente, **(B)** Bindungsfragmente, **(C)** funktionelle Gruppen und **(D)** höhere Strukturfragmente. Modifizierte Abbildung nach KOLSKÁ et al.^[41]

Wissensbasierte Methoden können jedoch auch zur Modellierung dreidimensionaler Molekülstrukturen, Protein/Protein- und Protein/Ligand-Komplexen herangezogen werden. Die wichtigsten Quellen für experimentell bestimmte Strukturinformationen sind die *Cambridge Crystallographic Database* (CSD) mit über 600.000 Kristallstrukturen kleiner organischer Moleküle sowie die *RCSB Protein Data Bank* (PDB) mit derzeit über 100.000 Protein- und DNA-Strukturen.^[46, 47]

Eine der häufigsten Aufgaben innerhalb des strukturbasierten Wirkstoffdesigns ist die Generierung plausibler Molekülkonformationen, also die Suche nach dem globalen Energieminimum. Man geht hierbei generell in drei Schritten vor. Zunächst werden plausible Startkonformationen generiert, anschließend wird von diesen Startpunkten aus eine Suche nach lokalen Minima durch Energieminimierung mittels Kraftfeld- oder quantenchemischer Methoden durchgeführt. In einem abschließenden Schritt werden die erhaltenen Strukturen verglichen, um Duplikate zu eliminieren. Da die Energieminimierung hierbei den mit Abstand rechenintensivsten Schritt darstellt, benötigt man aufgrund der konformationellen Vielfalt eines durchschnittlich großen Wirkstoffmoleküls stochastische oder wissensbasierte Methoden zur Ermittlung geeigneter Startkonformationen.^[48] Eine systematische Suche nach Startpunkten für eine Energieminimierung ist in den meisten Fällen nicht möglich: Bereits bei nur vier frei drehbaren Bindungen und einer systematischen Suche in 10°-Schritten ergäben

sich $36^4 = 1.679.616$ mögliche Startkonformationen, die energieminimiert werden müssten. In einem wissensbasierten Ansatz werden zunächst experimentell bestimmte Konformationen strukturell verwandter Moleküle oder Molekülfragmente, beispielsweise enthaltene Ringsysteme, analysiert. Der Algorithmus zur Generierung der Startkonformationen berücksichtigt anschließend nur diejenigen Geometrien, die mit den experimentellen Daten im Einklang stehen. Auf diese Weise werden nur sehr wenige Startkonformationen erzeugt, die zudem den proteingebundenen oder in Lösung stabilsten Konformationen am meisten ähneln.^[48]

Neben der Suche nach energetisch günstigen Konformationen und Bindungsmodi in aktiven Zentren von Rezeptoren ist deren energetische Bewertung, das sog. *Scoring*, ein Hauptaufgabengebiet des *Molecular Modeling*. Auch hier haben wissensbasierte Methoden in den letzten Jahren enorm an Bedeutung gewonnen. Gemein haben diese Methoden, dass aus experimentell bestimmten Kontaktgeometrien in Protein/Ligand-Komplexen statistische Potentiale berechnet werden, die angeben, wie häufig ein bestimmter Kontakt zwischen einer funktionellen Gruppe eines Liganden und einer Aminosäure im Rezeptor ist. Die daraus gewonnene Häufigkeitsverteilung lässt sich auf einen mittleren Referenzzustand beziehen, wodurch sich eine Energiefunktion berechnen lässt: Kontakte die häufiger auftreten als in der mittleren Verteilung sind demnach energetisch vorteilhaft, während seltenere Kontakte als energetisch ungünstig eingestuft werden.^[49, 50] Diese statistischen Potentiale können nicht nur zur Bewertung generierter Konformationen, sondern auch zur Analyse von Bindungstaschen in aktiven Zentren von Proteinen herangezogen werden und helfen, wissensbasierte Pharmakophormodelle aufzustellen (siehe Abschnitt 1.2.1).

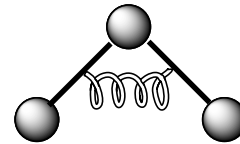
Kraftfeldmethoden: Ziel einer Kraftfeld- oder Molekülmechanik-Rechnung ist die Ermittlung einer energetisch günstigen dreidimensionalen Struktur eines molekularen Systems. Sie basiert dabei auf einer empirischen Kraftfeldgleichung mit anpassbaren Parametern (Abbildung 12). Die zentrale Annahme lautet: Bindungsabstände und -winkel nehmen näherungsweise Standardwerte an, es sei denn, sie werden durch anziehende oder abstoßende Wechselwirkungen davon abgehalten.^[40] Die Berechnungen sind der klassischen Mechanik entliehen und man betrachtet ein Molekül als Ensemble atomarer Kugeln. Diese werden, wie durch Federn verbundene Massenschwerpunkte, als harmonischer Oszillator behandelt.^[51]

$$E = E_{\text{Bindungslänge}} + E_{\text{Bindungswinkel}} + E_{\text{Torsion}} + E_{\text{nichtkovalent}}$$

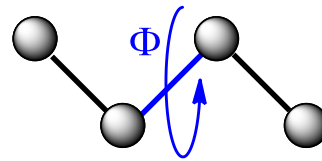
$$E = \frac{1}{2} \sum_{\text{Bindungslänge}} K_b (b - b_0)^2$$



$$+ \frac{1}{2} \sum_{\text{Bindungswinkel}} K_\theta (\theta - \theta_0)^2$$



$$+ \frac{1}{2} \sum_{\text{Torsionswinkel}} K_\Phi [1 + \cos(n\Phi - \delta)]^2$$



$$+ \sum_{\text{nichtkovalent}} \left(A_{ij} r_{ij}^{-12} - C_{ij} r_{ij}^{-6} + \frac{q_i q_j}{\epsilon r_{ij}} \right)$$

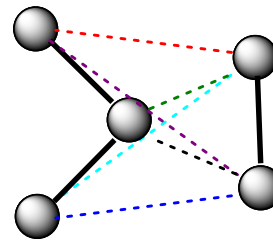


Abbildung 12: Eine typische Kraftfeldgleichung zur Berechnung der Gesamtenergie E eines molekularen Systems. Der erste Term beschreibt den Energiebeitrag der Dehnung oder Stauchung einer kovalenten Bindung mittels eines harmonischen Oszillators mit der Kraftkonstante K_b und der Gleichgewichtslänge b_0 als – experimentell oder rechnerisch – zu ermittelnde Parameter. Ebenfalls über ein harmonisches Potential wird eine Dehnung oder Stauchung des Bindungswinkels durch den zweiten Term erfasst. Parametrisiert werden müssen hierbei die Kraftkonstante K_θ und der Gleichgewichtswert θ_0 . Der dritte Term beschreibt die Änderung der Energie bei Änderung des Torsionswinkels und der letzte Term gibt den Einfluss nichtkovalenter Wechselwirkungen wieder. Hierbei werden die van-der-Waals-Kräfte meist durch das LENNARD-JONES-Potential mit dem Abstoßungsterm A_{ij}/r_{ij}^{12} sowie dem Anziehungsterm $-C_{ij}/r_{ij}^6$ berechnet. Elektrostatische Wechselwirkungen lassen sich durch das Coulomb-Potential $q_i q_j / \epsilon r_{ij}$ erfassen. Modifizierte Abbildung nach KLEBE.^[40]

Die Kräfte zwischen Punktladungen werden durch das Coulomb-Gesetz beschrieben, welches als kritische Parameter die Ladungen der betrachteten Teilchen und die relative Permittivität ϵ , auch Dielektrikum genannt, beinhaltet. Während die Bestimmung von Ladungen, beispielsweise mittels quantenmechanischer Berechnungen, sehr genau möglich ist, ist die Wahl des richtigen Dielektrikums für Kraftfeldrechnungen in Proteinen enorm schwierig und kann theoretisch alle Werte zwischen $\epsilon = 1$ für Vakuum und $\epsilon = 80$ für Wasser annehmen.^[40] Für Simulationen

von Proteinen sind sowohl feste Werte zwischen $\epsilon = 4$ und $\epsilon = 20$ gebräuchlich, als auch eine dynamische, umgebungsabhängige Parametrisierung. Generell werden Parametrisierungen von Kraftfeldern durch Angleichen an Ergebnisse möglichst genauer quantenchemischer Berechnungen und/oder an experimentelle Daten durchgeführt.^[52] Hierzu dienen beispielsweise röntgenkristallographisch gelöste 3D-Strukturen sowie Infrarot- und Ramanspektren, um Kraftkonstanten abzuleiten. Die hierbei gefundenen Parameter sind stark vom verwendeten Referenzdatensatz abhängig. Aus experimentellen Daten von Proteinen erhaltene Parameter können beispielsweise für die Modellierung von Zuckern oder Wirkstoffen völlig ungeeignet sein. Daher sind Kraftfelder, die ohne Parametrisierungen auskommen bzw. fehlende Parameter während der Berechnung selbst zuweisen, eine enorm nützliche Weiterentwicklung, welche u.a. in dem in dieser Arbeit häufig verwendeten Kraftfeld AMBER12:EHT eingesetzt wird.^[53]

Die NEWTON'sche Bewegungsgleichung erlaubt mithilfe des Energieerhaltungssatzes eine Simulation der Bewegungen von Atomen bzw. Molekülen unter der Einwirkung eines Kraftfelds. Für diese Moleküldynamiksimulationen (MD) werden mittels numerischer Verfahren die Bewegungsgleichungen für alle Teilchen gleichzeitig iterativ gelöst, wobei meist nur die zwischen zwei Teilchen wirkenden Kräfte gleichzeitig berücksichtigt werden (Paarnäherung).^[40] Auf diese Weise können dynamische Prozesse sichtbar gemacht werden, die in Zeitskalen von wenigen Nanosekunden ablaufen und experimentell nicht oder nur sehr aufwendig zu verfolgen wären. Eine weitere wichtige Anwendung ist die im vorhergehenden Abschnitt angesprochene Energieminimierung von molekularen Systemen zur Bestimmung energetisch günstiger Konformationen. In beiden Fällen werden zu Beginn der Moleküldynamiksimulation jedem Atom Startkoordinaten und -geschwindigkeiten zugewiesen, die – im Ensemblemittel – der gewünschten Temperatur entsprechen (Abbildung 13). Anschließend werden die Kräfte berechnet, die durch die umgebenden Teilchen auf das Atom wirken und es wird anhand der NEWTON'schen Bewegungsgleichung die neue Position des Atoms nach einem gewissen Zeitintervall (meist 1 fs) erhalten. Diese neue Position entspricht in der nächsten Iteration den Startkoordinaten.

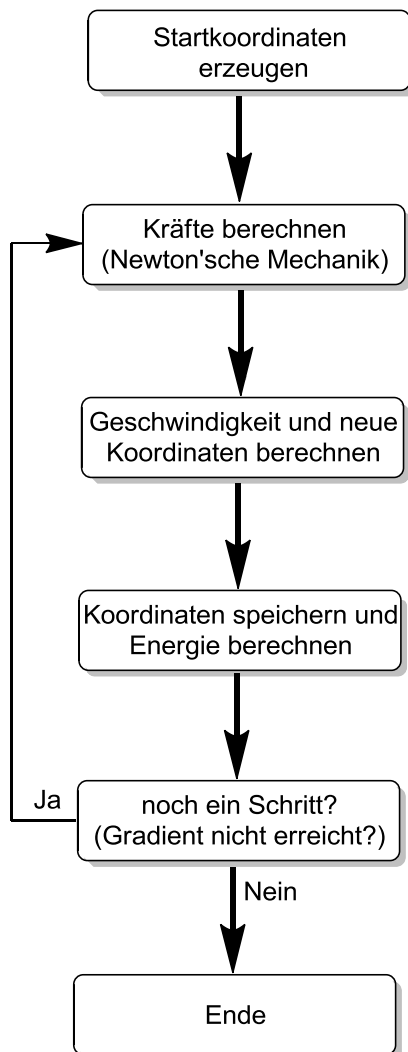


Abbildung 13: Schematischer Ablauf einer Moleküldynamikrechnung. Modifizierte Abbildung nach KLEBE.^[40]

Während man für die Simulation dynamischer Prozesse meist ein vorher festgelegtes Zeitintervall betrachtet, beispielsweise 50 ns, bricht man Energieminimierungen für gewöhnlich bei Erreichung eines bestimmten Energiegradienten ab. Hierbei wird der Energieunterschied zwischen zwei Iterationen anhand der Kraftfeldgleichung berechnet und mit einem zuvor festgelegten Wert verglichen.

1.2.1 Pharmakophor-Modellierung

Unter einem Pharmakophor (Ph4) bzw. einem Pharmakophormodell versteht man ganz allgemein die räumliche Anordnung von Interaktionsgruppen niedermolekularer Liganden, die für eine Bindung im aktiven Zentrum eines Proteins – und damit zur

Erzielung einer pharmakologischen Wirkung – erforderlich sind. Die einzelnen Interaktionsgruppen, wie beispielsweise wasserstoffbrückenbildende, hydrophobe, metallkomplexierende oder ionische funktionelle Gruppen, bezeichnet man häufig als Pharmakophordeskriptoren oder Pharmakophoreigenschaften.^[54]

Strukturbasierte Ph4-Modellierung: Ist die dreidimensionale Struktur eines aktiven Zentrums bekannt, so lässt sich hieraus ein Ph4 ableiten: Die Aminosäureseitenketten des Rezeptors lassen Rückschlüsse zu, wie funktionelle Gruppen des Liganden beschaffen sein müssen, um an sie zu binden. Bereits 1985 wurde von P. GOODFORD die Software GRID publiziert, mit der er genau diesen Ansatz verfolgte.^[55] Mittels eines Satzes niedermolekularer Sonden, beispielsweise einem Wassermolekül, einem aromatischen Kohlenstoff, einer H-Brücken-Donor- bzw. Akzeptorgruppe oder einer ionischen Gruppe, wird ein virtuell in das aktive Zentrum gelegtes Gitter abgerastert und die jeweilige Wechselwirkungsenergie für jeden Schnittpunkt dieses Gitters berechnet und gespeichert. Die Energien können anschließend graphisch dargestellt werden und deuten Brennpunkte für die jeweilige funktionelle Gruppe an. Zusammengenommen können die so erhaltenen Brennpunkte in ein Ph4-Modell übersetzt werden. Ein beeindruckendes Beispiel für den Erfolg dieses Ansatzes im rationalen Wirkstoffdesign ist die Entwicklung von Zanamivir, dem ersten Inhibitor der Influenza-Neuraminidase.^[56]

Auf die Möglichkeit, durch wissensbasierte statistische Potentiale bevorzugte Ligand/Protein-Wechselwirkungen und somit ein Ph4-Modell zu erhalten, wurde bereits in Abschnitt 1.2 detailliert eingegangen.

Die *Multiple Copy Simultaneous Search* (MCSS)-Methode funktioniert kraftfeldbasiert und bedient sich ebenfalls niedermolekularer Sonden.^[57] Hierbei werden jedoch kleine Moleküle (drei bis sechs Atome, möglichst wenig konformationelle Freiheitsgrade) eingesetzt. Wählt man diesen Satz an Molekülen geschickt, so lassen sich sterische und elektrostatische Wechselwirkungen realistisch simulieren und ein Großteil der organischen Verbindungen aus ihnen aufbauen. Hierzu gehören Fragmente wie Acetonitril, Acetat, Methanol, Dimethylether, Benzol oder Acetaldehyd. Viele Kopien dieser Sondenmoleküle werden im aktiven Zentrum positioniert und einer kraftfeldbasierten Energieminimierung unterzogen. Hierbei spüren die Sondenmoleküle nur die Wechselwirkung mit dem Protein, „sehen“ sich

aber untereinander nicht. Die Brennpunkte, an denen anschließend die meisten Sondenmoleküle zu finden sind, lassen sich direkt in ein Ph4-Modell übersetzen. Nachteil strukturbasierter Methoden ist im Allgemeinen, dass dynamische Konformationsänderungen des Rezeptors, ein sog. *induced-fit*, nicht oder nur sehr eingeschränkt berücksichtigt werden können.

Ligandenbasierte Ph4-Modellierung: Wird ein Ph4-Modell von einem Ensemble analog bindender Liganden abgeleitet statt von einer Proteinstruktur, so spricht man von einem ligandenbasierten Modell. 1979 wurde von G. H. MARSHALL et al. der sog. *active analog approach* vorgestellt.^[58] Hierbei werden zunächst jedem Molekül in einer Datenbank entsprechende Ph4-Deskriptoren zugewiesen, anschließend führt man für den ersten Eintrag der Datenbank eine detaillierte Konformationsanalyse durch. Gleichzeitig werden die Abstände der Ph4-Eigenschaften gespeichert und alle weiteren Moleküle der Datenbank werden auf diese Referenzgeometrie bezogen. Besitzen die Moleküle unterschiedliche konformationelle Freiheitsgrade in verschiedenen Bereichen ihrer Gerüste, so bleiben am Ende nur sehr wenige Ph4-Hypothesen übrig, die von allen Verbindungen der Datenbank erfüllt werden. Entscheidend für die Effizienz dieser Methode sind die Wahl des richtigen Startmoleküls sowie das Auffinden von Konformationen, die dem gebundenen Zustand möglichst exakt entsprechen. Die in dieser Arbeit verwendete Methode (siehe Abschnitt 3.5) unterscheidet sich insofern vom *active analog approach*, als dass diejenigen Ph4-Hypothesen gesucht werden, die von möglichst vielen aktiven Verbindungen erfüllt werden. Die Reihenfolge der Moleküle in der Eingangsdatenbank ist somit unerheblich. Da die Zahl an möglichen Ph4-Modellen aktiver Verbindungen, je nach Größe des betrachteten konformationellen Raumes, gegen unendlich geht, wird ein baumartiger Algorithmus angewendet, der größere Pharmakophore aus kleineren ableitet (Abbildung 14). Auf diese Weise verhindert man eine drohende kombinatorische Explosion und erhält dennoch einen ausreichend vollständigen Satz an komplexen Ph4-Hypothesen.

Neben der Erzeugung von Ph4-Modellen spielt deren anschließende Bewertung eine entscheidende Rolle. Eine Möglichkeit des *Scorings* beruht auf der Annahme, dass sich alle aktiven Moleküle möglichst gut an einem plausiblen Ph4 ausrichten lassen müssen. Eine solche Ausrichtung und deren anschließende Bewertung kann

beispielsweise mit der Software FlexAlign von P. LABUTE et al. durchgeführt werden.^[59]

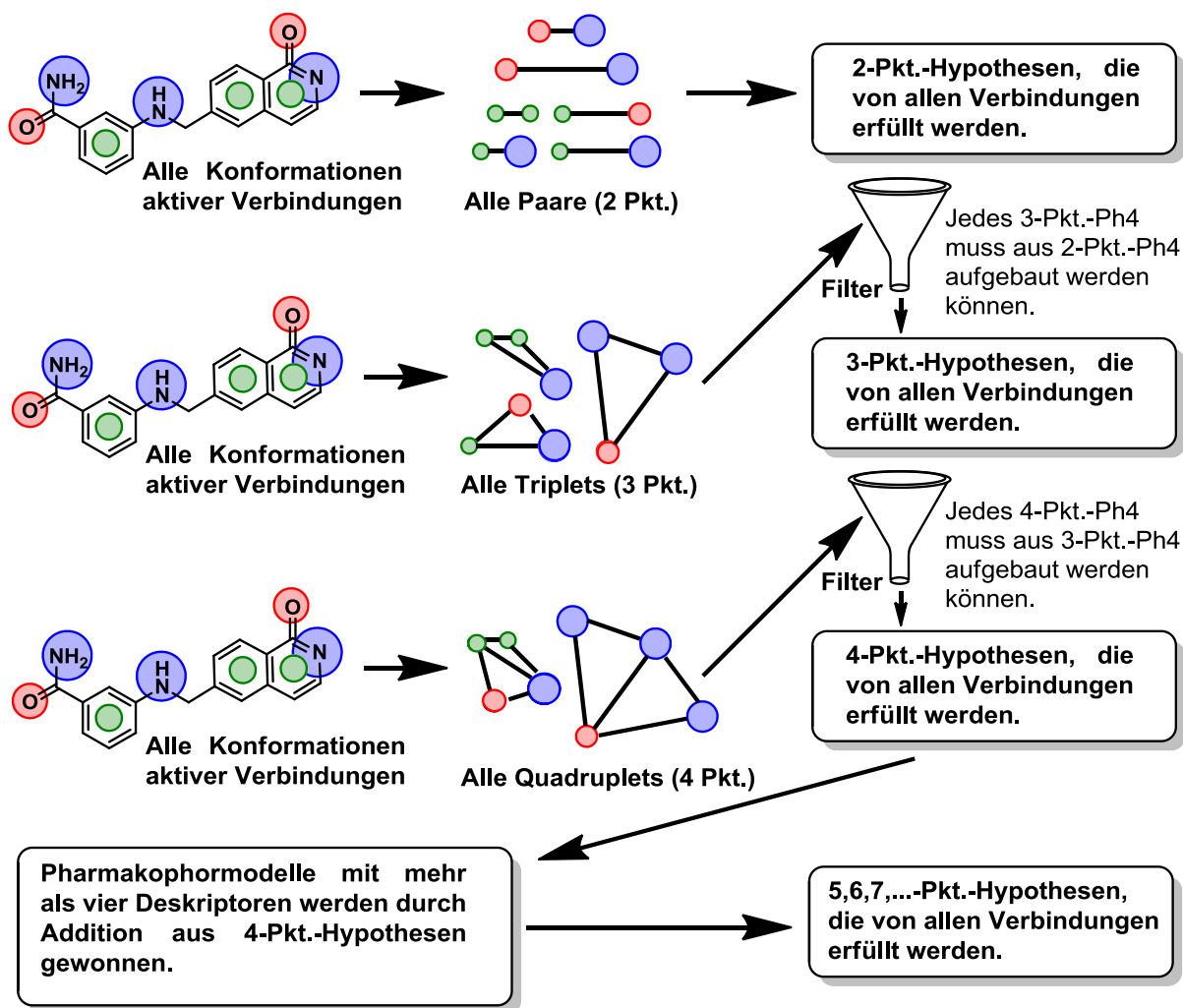


Abbildung 14: Baumartiger Algorithmus für die systematische Suche nach komplexen Pharmakophorhypothesen. Ausgehend von einer Datenbank analog bindender, aktiver Moleküle (dem sog. Trainings-Set) wird eine detaillierte Konformationssuche durchgeführt. Für jede plausible Konformation werden anschließend alle Deskriptorpaare gesucht und als 2-Pkt.-Hypothesen abgespeichert. Im zweiten Schritt werden alle 3-Pkt.-Hypothesen gesucht, jedoch nur diejenigen abgespeichert, die aus den bekannten 2-Pkt.-Hypothesen aufgebaut werden können. Nach 4-Pkt.-Hypothesen wird ebenfalls systematisch gesucht, wobei diese die bekannten 3-Pkt.-Hypothesen beinhalten müssen. Komplexere Pharmakophormodelle mit fünf oder mehr Deskriptoren werden aus diesem Satz kleinerer Pharmakophore erstellt. Nach jedem Schritt werden die erhaltenen Pharmakophormodelle gegen das Trainings-Set validiert und nur diejenigen im Folgeschritt verwendet, die von allen aktiven Verbindungen erfüllt werden. Modifizierte Abbildung nach einer Vorlage aus dem Handbuch der Software MOE2014.09.^[60]

Häufig werden Ph4-Modelle genutzt, um Datenbanken schnell und effizient nach neuen Leitstrukturen zu durchsuchen. Für eine solche Anwendung ist die Fähigkeit des Modells, mutmaßlich aktive von mutmaßlich inaktiven Verbindungen zu

unterscheiden, essentiell. Um diese Treffsicherheit der Ph4-Hypothesen zu überprüfen, validiert man sie gegen ein Set aus bekanntermaßen aktiven und bekanntermaßen inaktiven Verbindungen mit ähnlicher Molekülstruktur.

1.2.2 Quantitative Struktur/Aktivitäts-Beziehung (QSAR)

Unter einer Struktur/Aktivitäts-Beziehung (SAR) versteht man nach IUPAC die Abhängigkeit der biologischen Aktivität eines Moleküls von dessen dreidimensionaler Struktur.^[61] Die Analyse der SAR ermöglicht demnach die Identifizierung struktureller Merkmale eines Ensembles von Molekülen, die verantwortlich für eine bestimmte Wirkung in einem Organismus sind. Eine Möglichkeit der SAR-Analyse bietet die in Abschnitt 1.2.1 vorgestellte Pharmakophor-Modellierung. Die zugrunde liegende Annahme aller Formen der SAR-Analyse ist, dass ähnliche Moleküle eine ähnliche Aktivität besitzen müssen. Die Definition dieser „Ähnlichkeit“ ist jedoch nicht trivial. Sich im Hinblick auf Struktur und Pharmakophordeskriptoren ähnelnde Moleküle können dennoch dramatisch unterschiedliche biologische Aktivitäten besitzen. Beispielhaft hierfür sind die in Abschnitt 1.1.4 erwähnten Proteasominhibitoren Bortezomib **5** und dessen Aldehydanalogon **6**. Sie unterscheiden sich lediglich in ihrer elektrophilen Kopfgruppe und weisen dennoch um mehrere Zehnerpotenzen verschiedene IC₅₀-Werte gegenüber der β 5-Untereinheit auf. Um Wirkstoffe hinsichtlich metabolischer, toxikologischer und verschiedenster physikochemischer Eigenschaften zu optimieren, ohne dabei deren biologische Aktivität zu verlieren, ist daher die Kenntnis von Molekülfragmenten und funktionellen Gruppen entscheidend, die eine sogenannte Bioisosterie aufweisen – in biologischen Systemen also eine ähnliche Wirkung zeigen.^[62, 63]

Die quantitative Beschreibung der SAR mittels eines mathematischen Modells wird quantitative Struktur/Aktivitäts-Beziehung (QSAR) genannt und gibt die Aktivität gegenüber einem bestimmten Zielprotein als eine Funktion physikochemischer oder struktureller Deskriptoren wieder. Die analysierten Substanzen sollten hierbei aus einer chemisch einheitlichen Serie stammen und ihre Aktivität muss durch Modulation eines identischen biologischen Zielmoleküls sowie einem ähnlichen Wirkmechanismus herrühren. Setzt man den Begriff „Aktivität“ mit dem Begriff „Reaktivität“ gleich, so lassen sich mit Hilfe von QSAR-Methoden mathematische Modelle zur Vorhersage

verschiedenster physikochemischer Eigenschaften aus Molekülstrukturen ableiten. Bereits 1937 publizierte L. P. HAMMETT eine Beziehung zwischen den elektronischen Eigenschaften von Substituenten und der Reaktivität aromatischer Verbindungen.^[64] Durch experimentelle Bestimmung der Dissoziationskonstanten substituierter Benzoessäurederivate war er in der Lage, die sog. HAMMETT-Gleichung (Gleichung 1) aufzustellen:

$$\rho * \sigma = \log \left(\frac{K_X}{K_H} \right)$$

Gleichung 1: HAMMETT-Gleichung. ρ = Reaktionskonstante, σ = Hammett-Konstante, K_X = Gleichgewichtskonstante für substituierten Aromaten, K_H = Gleichgewichtskonstante für unsubstituierten Aromaten.^[64]

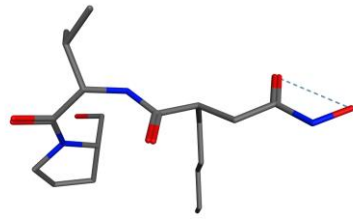
Während die HAMMETT-Konstante σ unabhängig von der betrachteten Reaktion ist und nur von der elektronischen Struktur des Substituenten X abhängt, muss die Reaktionskonstante ρ für jede Reaktionsart experimentell bestimmt und angepasst werden. Auf diese Weise konnte die Allgemeingültigkeit der HAMMETT-Gleichung anhand von 38 verschiedenen Reaktionen bereits in der Originalpublikation verifiziert werden.^[64]

Erwähnenswert ist an dieser Stelle, dass ein Verhältnis von Gleichgewichtskonstanten einem Verhältnis von freien Reaktionsenthalpien äquivalent ist. Die HAMMETT-Gleichung ist somit eine der ersten und vermutlich die bekannteste lineare Freie-Enthalpie-Beziehung (LFER). Insbesondere die Beeinflussung der Elektronendichte eines Heteroatoms an einem aromatischen System kann dessen Fähigkeit, H-Brücken auszubilden, deutlich verändern, weshalb die Anwendung der HAMMETT-Konstante σ als elektronischer Parameter in QSAR-Modellen äußerst gebräuchlich ist. Mit Hilfe quantenchemischer *ab initio*-Berechnungen konnten die zunächst nur empirisch ermittelten Zusammenhänge zudem theoretisch untermauert werden.^[65]

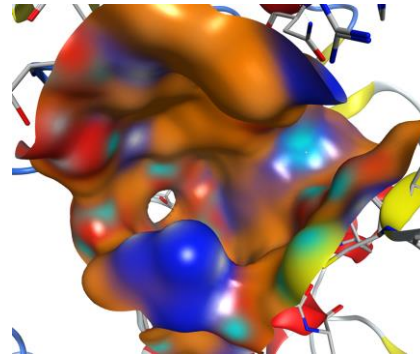
1.2.3 Kovalentes Docking

Durch molekulares Docking, genauer Protein/Ligand-Docking, werden niedermolekulare Wirkstoffkandidaten computerbasiert in die Bindungstasche eines Zielproteins eingepasst (Abbildung 15). Ausgangspunkt dieser strukturbasierten Methode ist eine experimentell bestimmte oder modellierte dreidimensionale Struktur

A

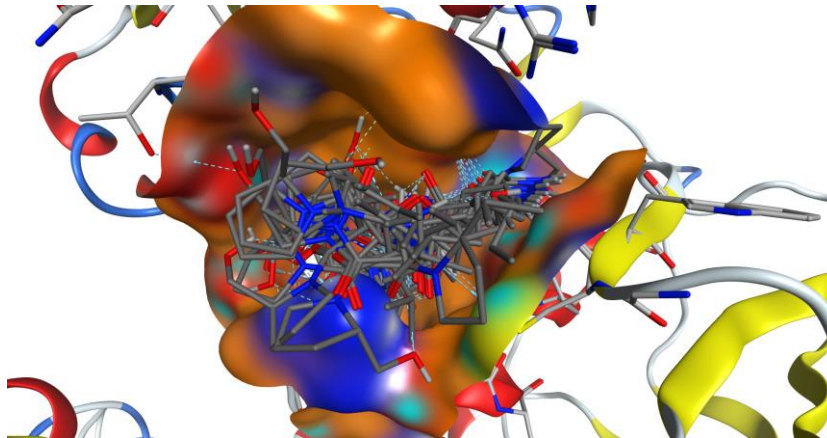


Konformationsdatenbank
des Liganden



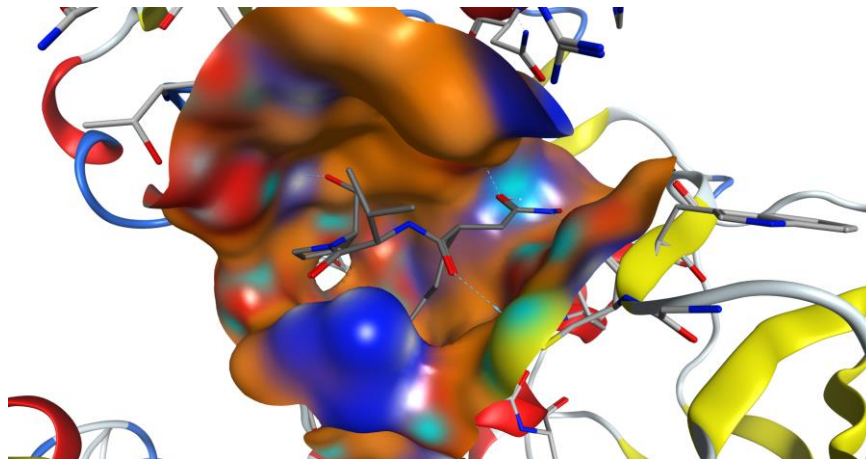
3D-Struktur des Rezeptorproteins

B



Durch geschickte Platzierung der Ligandenkonformationen in das aktive Zentrum werden verschiedene Bindungshypothesen erhalten.

C



Mit Hilfe von *Scoring*-Funktionen und ähnlichen Bewertungsmethoden wird die plausibelste der generierten Hypothesen identifiziert.

Abbildung 15: Schema einer typischen Dockingsimulation. **(A)** Benötigt werden die 3D-Struktur des Rezeptorproteins und ein Ensemble von Ligandenkonformationen, welches entweder während des Dockings erzeugt oder durch eine detaillierte Konformationsanalyse erhalten wurde. **(B)** Während des sog. *Placements* werden die Ligandenkonformationen mittels spezieller Algorithmen in dem aktiven Zentrum platziert. Dies kann auch anhand vorher erstellter Pharmakophormodelle oder wissensbasierter Häufigkeitsverteilungen geschehen. **(C)** Die Identifizierung eines möglichst plausiblen (*near native*) Bindungsmodus erfolgt meist mittels kraftfeldbasierter, wissensbasierter oder empirischer *Scoring*-Funktionen.

des aktiven Zentrums sowie ein Konformationsensemble des Liganden.^[66, 67] Da der konformationelle Raum innerhalb eines aktiven Zentrums für gewöhnlich sehr eingeschränkt ist, bieten sich für eine Konformationsanalyse hierbei besonders wissens- oder pharmakophorbasierte Ansätze an. Die Dockingsoftware ermittelt anschließend plausible Bindungsgeometrien, indem sie die vorgegebenen Konformationen anhand geeigneter Algorithmen in der Bindungstasche platziert (engl.: *placement*). Ist für das aktive Zentrum bzw. die jeweilige Ligandenklasse ein Pharmakophormodell bekannt, so sollte dieses möglichst schon während des *Placements*, spätestens aber zum Eliminieren ungünstiger Bindungsgeometrien, in die Berechnungen einfließen. Essenziell für den Erfolg einer Dockingsimulation ist die korrekte Bewertung der erzeugten Bindungshypothesen. Es gilt, mittels geeigneter Verfahren, aus einer Vielzahl geometrisch möglicher Bindungsmodi diejenigen herauszufiltern, die eine plausible, dem Experiment entsprechende Lösung abgeben. Dockingstudien werden für gewöhnlich an sehr vielen Liganden durchgeführt; verwendete *Scoring*-Funktionen müssen also nicht nur möglichst akkurate, sondern vor allem schnelle Abschätzungen für die freie Enthalpie ΔG der Ligandenbindung liefern. Regressionsbasierte *Scoring*-Funktionen basieren auf einer Linearkombination von Enthalpie- und Entropiebeiträgen, die anhand eines Referenzdatensatzes empirisch ermittelt wurden. Das Vorgehen erinnert sehr stark an das Aufstellen einer QSAR-Gleichung (siehe Abschnitt 1.2.2). Beispielhaft hierfür ist die in dieser Arbeit häufig verwendete *Scoring*-Funktion *London dG* (Gleichung 2).

$$\Delta G = c + E_{\text{flex}} + \sum_{\text{H-Brücken}} c_{\text{HB}} f_{\text{HB}} + \sum_{\text{Metall}} c_{\text{M}} f_{\text{M}} + \sum_{\text{Desolv.}} \Delta D$$

Gleichung 2: Die *Scoring*-Funktion *London dG*. ΔG = Freie Bindungsenthalpie, c = Änderung der Rotations-/Translationsentropie, E_{flex} = Energieänderung durch Flexibilitätsverlust, $c_{\text{HB}} f_{\text{HB}}$ = Energieänderung durch H-Brücken, $c_{\text{M}} f_{\text{M}}$ = Energieänderung durch Ligand/Metall-Kontakte (falls vorhanden), ΔD = Energieänderung durch Desolvatationseffekte.^[60]

Sie besteht aus einer Summe von Termen, wobei c für den Gewinn/Verlust an Rotations-/Translationsentropie steht, E_{flex} für die Energieänderung aufgrund von Flexibilitätsverlusten, c_{HB} für den Energiegewinn durch H-Brücken (mit dem Faktor f_{HB} zur Erfassung von Abweichungen vom idealen Verhalten), c_{M} für den Energiegewinn durch eventuelle Ligand/Metall-Wechselwirkungen (mit dem Faktor f_{M} zur Erfassung

von Abweichungen vom idealen Verhalten) und ΔD für die Änderung der freien Enthalpie aufgrund von Desolvatationseffekten.^[60] Kraftfeldbasierte *Scoring*-Funktionen berechnen Enthalpiebeiträge über eine Kraftfeldgleichung und nähern entropische Anteile an der freien Bindungsenthalpie meist über einen Zusatzterm an. Für die Erfassung von Entropiebeiträgen zur freien Bindungsenthalpie, insbesondere den Entropiegewinn durch Verdrängung gebundenen Wassers aus dem aktiven Zentrum, ist bis heute noch keine wirklich überzeugende Lösung gefunden worden.^[40] Aus diesem Grund liefern wissensbasierte *Scoring*-Funktionen oft wesentlich genauere Ergebnisse im Hinblick auf korrekt vorhergesagte Bindungsmodi. Sie sind nicht auf eine explizite Berechnung der freien Bindungsenthalpie angewiesen: Die ihnen zugrunde liegenden statistischen Potentiale entsprechen bereits Bindungsgeometrien realer Protein/Ligand-Komplexe.^[49]

Heute stehen für Dockingsimulationen eine Vielzahl kommerzieller und kostenloser Softwarelösungen zur Verfügung.^[66] Die Zahl derer, die mit einer kovalenten Bindung zwischen Ligand und Rezeptor umgehen können, ist jedoch weitaus geringer.^[68] Dies ist zum einen auf ein in der Vergangenheit eher geringes Interesse an kovalent-bindenden Wirkstoffen zurückzuführen, welches erst in den letzten Jahren deutlich gestiegen ist.^[69] Zum anderen gibt es technische Unterschiede bezüglich der Ligandenvorbereitung, des *Placements* und des *Scorings*, für die sich verhältnismäßig langsam neue Methoden etablieren. Die Ligandenvorbereitung gestaltet sich insofern schwieriger, als dass sich aufgrund einer kovalenten Bindung der nukleophilen Seitenkette des Rezeptors mit der elektrophilen Kopfgruppe eines Liganden deren Konstitution und Geometrie ändert. Zahlreiche Beispiele hierfür finden sich in Abschnitt 1.1.4. Um ein kovalentes Docking automatisiert durchführen zu können, muss die Software die elektrophile Kopfgruppe eines Liganden erkennen und sie daraufhin über den entsprechenden Bindungsmechanismus in die korrespondierende gebundene Geometrie überführen. Viele elektrophile Kopfgruppen sind zudem prochiral – durch den Angriff des Nukleophils bilden sich also häufig zwei oder mehr Stereoisomere –, was ebenfalls berücksichtigt werden muss. Für die Platzierung eines kovalent gebundenen Liganden im aktiven Zentrum gibt es bereits mehrere literaturbekannte Ansätze. In der Software GOLD beispielsweise wird ein Ankeratom definiert, das sowohl Teil des Liganden, als auch Teil des Rezeptors ist und dessen Position während des Dockings fixiert bleibt.^[70] Für diese Software wurde von SCHRÖDER et al.

eine automatische Ligandenvorbereitung entwickelt, was die Methode – als eine der wenigen – für virtuelle Screenings größerer Ligandenbibliotheken nutzbar macht.^[71] Die beiden aktuellsten Methoden sind CovDock^[72], welches auf der Softwareplattform Schrödinger beruht, und DOCKkovalent^[73], welches den Dockingalgorithmus von DOCK3.6^[74] nutzt. Das Besondere an CovDock ist die realitätsnahe Grundhypothese, dass vor einer kovalenten Reaktion zunächst ein energetisch günstiger ungebundener Zustand vorliegen muss. Diese Hypothese erhält Einzug in den Dockingalgorithmus, indem zunächst ein klassisches nichtkovalentes Docking durchgeführt wird. Um eine sterische Abstoßung zu vermeiden, wird hierbei die nukleophile Seitenkette gegen Alanin ausgetauscht. Erst anschließend wird die kovalente Bindung virtuell gebildet und die vorher erzeugten Bindungsmodi werden optimiert. Die zusätzliche Annahme, dass sich die ungebundenen Bindungsgeometrien durch Ausbildung der kovalenten Bindung nur marginal verändern, ist jedoch in vielen Fällen die Schwachstelle von CovDock. Kommt es hierbei beispielsweise zu einer Unterbrechung konjugierter Systeme, einer Ringöffnung oder Eliminierung einer Abgangsgruppe, so führt diese Annahme zu absehbaren Komplikationen.

Eines der wenigen Programme, das versucht, die Bindungsenergie der kovalenten Bindung explizit anzunähern, ist CovalentDock.^[75] Die Entwickler nutzen hierfür das MORSE-Potential mit quantenchemisch berechneten Parametern für die beiden untersuchten Bindungsmechanismen: Michael-Akzeptor bindet an Cystein und β -Lactam bindet an Serin.^[76] Elektronische Einflüsse des Liganden auf die Elektrophilie seiner Kopfgruppe bleiben unberücksichtigt. Hier werden die Probleme des *Scorings* kovalenter Liganden deutlich. Der energetische Beitrag der kovalenten Bindungsbildung ist so enorm, dass man ihn nicht vernachlässigen sollte. Andererseits ist es bei einem dockingbasierten virtuellen *Screening* wichtig, Ligandenbibliotheken mit einer großen Bandbreite elektrophiler Kopfgruppen erfassen zu können, was eine vorherige Parametrisierung erschwert. Zudem hängt die Elektrophilie der Kopfgruppe vom Rest des Ligandenmoleküls ab, was nur durch zeitintensive quantenchemische Rechnungen zu erfassen wäre. Diese Problematiken sind bislang nicht zufriedenstellend gelöst worden, weshalb die meisten Programme den Energiebeitrag der kovalenten Bindung schlicht ignorieren oder ihn durch einen Standardwert ersetzen. Bei der Interpretation der Dockingergebnisse, insbesondere beim Vergleich

von Liganden mit verschiedenen Kopfgruppen, muss dieser Umstand berücksichtigt werden.

1.3 Die ALZHEIMER-Demenz

„Ich beginne jetzt eine Reise, die mich in den Sonnenuntergang meines Lebens führt.“

– Mit diesen Worten wandte sich der ehemalige US-Präsident und frühere Schauspieler R. REAGAN 1994 an die amerikanische Nation, um die kürzlich bei ihm diagnostizierte Demenzerkrankung Morbus ALZHEIMER ins Licht der Öffentlichkeit zu bringen.^[77] Auf dieser beängstigenden Reise verlieren die Betroffenen nach und nach ihr Gedächtnis und jegliche kognitive Fähigkeiten. Gemütsschwankungen bis hin zum vollkommenen Schwinden der eigenen Persönlichkeit charakterisieren den Verlauf dieser unheilbaren Krankheit. Durch den fortschreitenden Verlust an Nervenzellen vermindert sich das Volumen der Gehirnregionen, die für die Informationsverarbeitung und das Gedächtnis verantwortlich sind (Gehirn-Atrophie).

1906 beschrieb der Psychiater und Neuropathologe A. ALZHEIMER in seiner Schrift „Über eine eigenartige Erkrankung der Hirnrinde“ erstmals den Verlauf der Krankheit anhand seiner Patientin A. DETER. Heute, über ein Jahrhundert später, ist die ALZHEIMER-Demenz (AD) eines der wichtigsten Themen neurowissenschaftlicher Forschung. Mit zunehmender Lebenserwartung steigt das Risiko, an einer Demenz wie Morbus ALZHEIMER zu erkranken. Bei den 40- bis 64-Jährigen sind Demenzen sehr selten; die Erkrankungsrate liegt hier bei weniger als 0,1 Prozent. Das entspricht rund 20.000 Menschen in Deutschland. Ab diesem Alter steigt die Erkrankungsrate steil an. Als Faustregel kann gelten: Alle fünf Lebensjahre verdoppelt sich die Prävalenz. Sie liegt bei den 65- bis 69-Jährigen bei rund 1% und klettert bei den über 90-Jährigen, je nach Studie, auf 28 bis 40%.^[78]

Die eigentliche Ursache der AD ist trotz aller Fortschritte in der Forschung noch immer unbekannt.^[79] Es wird vermutet, dass die charakteristischen Proteinablagerungen im Gehirn für die Zerstörung des neuronalen Gewebes ursächlich verantwortlich sind. Es wird dabei zwischen den intrazellulären neurofibrillären Bündeln (NFTs) und den extrazellulären amyloiden Plaques (APs) unterschieden, welche damals wie heute die Grundlage für die histologische *post mortem*-Diagnose der AD bilden.^[80]

1.3.1 Die Amyloid-Hypothese

Die Akkumulation der APs im Gehirn beginnt schon 10-15 Jahre vor dem Auftreten erster Symptome und ist kein verlässlicher Marker dafür, ob ein Mensch im Laufe seines Lebens an Morbus ALZHEIMER erkrankt.^[81] Dennoch wird die Amyloid-Hypothese, nach der die Bildung von APs als ursächlich für die Entwicklung von AD betrachtet wird, von der Mehrheit der Experten unterstützt. Die Ausbreitung der APs in verschiedenen Gehirnregionen verläuft über vier Phasen.^[82] Zunächst treten sie im temporalen basalen Neocortex auf und breiten sich später auf die entorhinale Region und den mittleren Temporallappen aus, bis sie schließlich auch im Hippocampus, im Hirnstamm und im Kleinhirn zu finden sind.

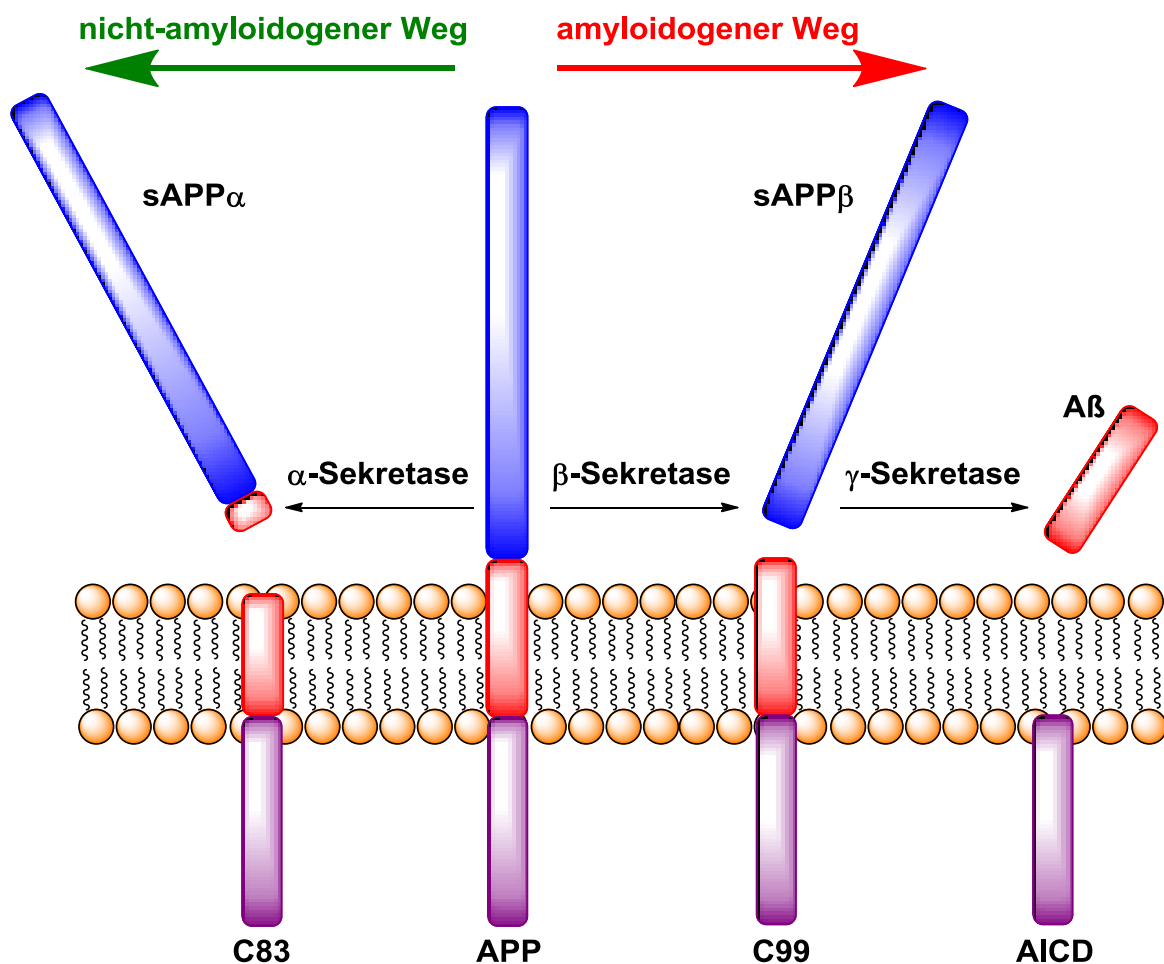


Abbildung 16: Schematische Abbildung der APP-Prozessierung. Auf dem nicht amyloidogenen Weg (grün) wird APP zunächst von der α -Sekretase geschnitten, woraufhin das N-terminale Fragment sAPP α abgespalten wird. Auf dem amyloidogenen Weg (rot) wird APP von der β -Sekretase in das N-terminale Fragment sAPP β und den C-terminalen Transmembranteil C99 gespalten. Dieser wiederum wird von der γ -Sekretase zu AICD und dem A β -Peptid umgesetzt. Modifizierte Abbildung nach LAFERLA et al.^[83]

Die APs bestehen hauptsächlich aus dem Amyloid- β -Peptid ($A\beta$). $A\beta$ wird aus dem Amyloid-Vorläufer-Protein (APP) durch das Zusammenspiel verschiedener Aspartylproteasen gebildet. Bei gesunden Menschen liegt der Anteil der „gutartigen“ nicht-amyloidogenen APP-Prozessierung bei über 90%. Die amyloidogene Prozessierung, welche auf einen gestörten APP-Metabolismus zurückzuführen ist, setzt sich aus zwei aufeinanderfolgenden Schritten zusammen (Abbildung 16). APP wird zunächst von der β -Sekretase geschnitten. Dabei entstehen auf der Seite des N-Terminus des APP ein lösliches Fragment namens sAPP β und ein membrangebundenes 99-Aminosäuren-langes Fragment C99 auf der Seite des C-Terminus. Die Spaltung durch die γ -Sekretase bildet nun unterschiedlich lange $A\beta$ -Peptide, von denen das 42-Aminosäuren-lange $A\beta_{42}$ am effektivsten aggregiert und somit am toxischsten ist.^[83]

1.3.2 Die Tau-Hypothese

Das Auftreten der intrazellulären NFTs korreliert wesentlich besser mit den ersten kognitiven Symptomen der Krankheit, als es beim Vorhandensein von APs der Fall ist.^[81] NFTs bestehen aus einer hyperphosphorylierten Form des Tau-Proteins, dessen sechs Isoformen in gesunden Neuronen für die Stabilisierung axonaler Mikrotubuli durch Interaktion mit Tubulin verantwortlich sind.^[84]

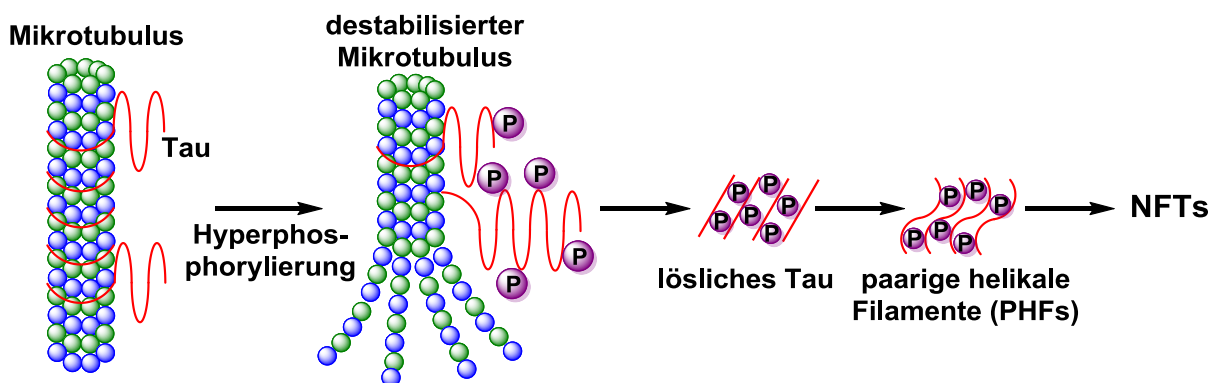


Abbildung 17: Schematischer Mechanismus der NFT-Bildung. Durch Hyperphosphorylierung verliert das Tau-Protein seine Affinität zu den Mikrotubuli und ist nicht weiter in der Lage sie zu stabilisieren. Dies führt zu ihrer Depolymerisation. Das Tau-Protein liegt zunächst als hyperphosphoryliertes lösliches Tau vor und bildet aufgrund seiner Agglomerationsneigung in einem ersten Schritt paarige helikale Filamente (PHFs), welche im Anschluss weiter koagulieren und die größeren unlöslichen Neurofibrillenbündel (NFTs) bilden. Modifizierte Abbildung nach MAZANETZ et al.^[85]

Eine Vielzahl an Kinasen wurde bisher mit der Hyperphosphorylierung von Tau in Verbindung gebracht, darunter die Glykogen-Synthase-Kinase 3 β (GSK3 β), die cyclin-abhängige Kinase 5 (CDK5), die extrazelluläre signalregulierte Kinase 2 (ERK2) und die *dual specificity tyrosine-phosphorylation-regulated kinase 1A* (DYRK1A).^[85, 86] Durch die Hyperphosphorylierung kommt es durch Störung der Wechselwirkungen zwischen Tau und den Mikrotubuli zu einer Ablösung von zunächst löslichen Tau-Proteinen, wodurch die Mikrotubuli destabilisiert werden und teilweise depolymerisieren (Abbildung 17). Hyperphosphoryliertes Tau neigt stark zur Selbstagglomeration, was zunächst zu Ablagerungen in Form paariger helikaler Filamente (PHFs) und im weiteren Verlauf zu den größeren unlöslichen NFTs führt. Der Verlust der Mikrotubuli zusammen mit der Ablagerung großer Mengen NFTs im Gehirn führt zu einem gestörten axonalen Transport und somit zum Absterben der Nervenzellen.^[85]

1.3.3 Diagnose durch Fluoreszenz-Imaging

Ebenso essentiell wie die Entwicklung von Therapiemöglichkeiten ist die Entwicklung von kostengünstigen, einfach durchzuführenden und zuverlässigen Diagnoseverfahren für die AD in einem möglichst frühen Stadium des Krankheitsverlaufs. Es ist dabei entscheidend herauszufinden, ob eine etwaige Gedächtnisstörung beispielsweise auf einen Tumor, eine Durchblutungsstörung oder eine Art der Demenz zurückzuführen ist. In den letzten Jahren wurden Ärzten mit neuen psychometrischen Verfahren effektive Werkzeuge an die Hand gegeben. Beispielsweise der von Wissenschaftlern am Max-Planck-Institut für neurologische Forschung in Köln entwickelte *Rapid Dementia Screening Test* (RDST).^[87] Er besteht aus zwei Aufgaben: Zum einen aus einer "Wortgenerierungsaufgabe", bei der in einer Minute möglichst viele Wörter zu einem vorgegebenen Oberbegriff genannt werden müssen, und zum anderen aus einer "Zahlentranskodierungsaufgabe", bei der arabische Zahlen in Zahlwörter – wie beim Schreiben eines Schecks – und umgekehrt Zahlwörter in arabische Zahlen umgewandelt werden müssen. In einer Studie mit 490 Probanden konnten mit dem RDST von den Patienten mit leichter bis mittelgradiger Demenz 80% korrekt klassifiziert werden, während es mit herkömmlichen Klassifizierungen nach ICD-10 nur etwa 25% sind.

Unglücklicherweise geht jedoch den kognitiven Störungen und somit den ersten Anzeichen einer Demenz ein oft schon jahrelanger neuronaler Verfall voraus. Deshalb ist es von größter Wichtigkeit, Ablagerungen von A β und NFTs im Gehirn zu detektieren, zu lokalisieren und zu quantifizieren. Dadurch ließen sich schon Stadien der AD erkennen, bei denen noch keinerlei oder höchstens sehr leichte Beeinträchtigungen der kognitiven Fähigkeiten bemerkbar sind. Darüber hinaus würde eine genauere Diagnostik bei der Entwicklung von Therapien helfen, indem man deren Erfolg pathologisch verfolgen könnte. Substanzen, die selektiv an APs bzw. NFTs binden, liegen deswegen im Fokus wissenschaftlicher Arbeiten. Vor allem zwei nicht invasive bildgebende Verfahren wurden in den letzten Jahren zur Lokalisation und Quantifizierung dieser Proteinaggregate genutzt: Die Positronen-Emissions-Tomographie (PET) und die Einzelphotonen-Emissions-Computertomographie (SPECT).^[88, 89] Beiden Verfahren liegt zugrunde, dass sie radioaktiv markierte Substanzen im Gewebe sichtbar machen. Jedoch sind sie apparativ aufwendig, teuer und wegen der Strahlenbelastung für eine vorsorgliche, verdachtsunabhängige Diagnose ungeeignet.

Große Hoffnung wird auf die Anwendung der Fluoreszenz- und Zwei-Photonen-Fluoreszenz-Mikroskopie als bildgebende Methoden in der AD-Diagnostik gesetzt. Sowohl AP- als auch NFT-Ablagerungen konnten in der murinen und menschlichen Retina nachgewiesen werden, was völlig neue Möglichkeiten der optischen Bildgebung ermöglicht.^[90-92] Bei der Fluoreszenz handelt es sich um einen photophysikalischen Prozess, bei dem der Fluorophor Energie in Form elektromagnetischer Strahlung absorbiert und dadurch in einen elektronisch angeregten Zustand überführt wird (Abbildung 18A). Im Falle der Zwei-Photonen-Fluoreszenz wird dies durch gleichzeitige Anregung mit zwei Photonen halber Energie erreicht. Anschließend wird ein Teil der Anregungsenergie strahlungsfrei an die Umgebung abgegeben, bis der Schwingungsgrundzustand des elektronisch angeregten Zustands erreicht ist. Die Rückkehr in den elektronischen Grundzustand erfolgt durch Emission der frei werdenden Energie in Form elektromagnetischer Strahlung höherer Wellenlänge. Zu Ehren des Entdeckers der Fluoreszenz – G. STOKES – nennt man die Wellenlängendifferenz zwischen Anregung und Emission auch STOKES-Verschiebung.^[93]

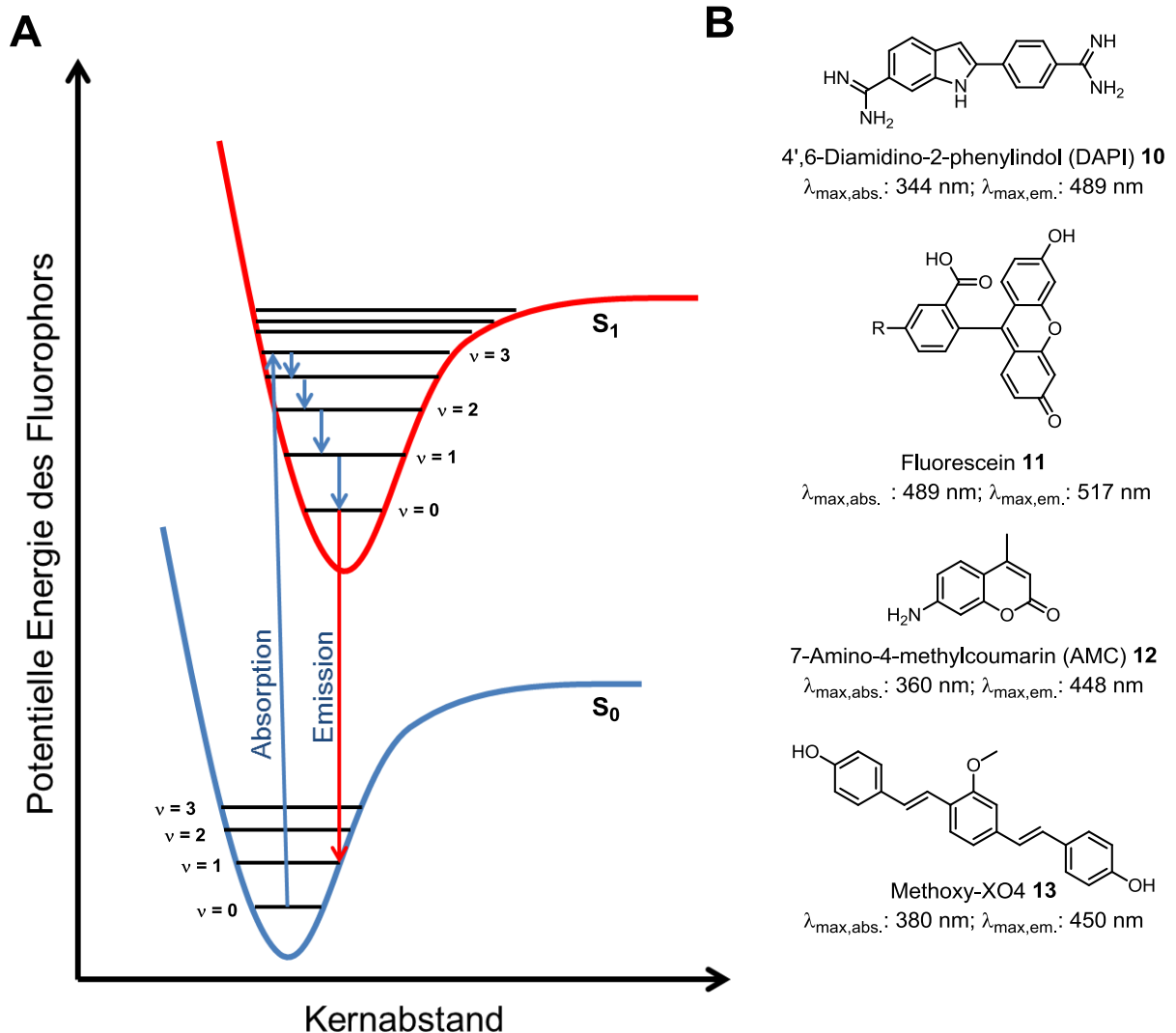


Abbildung 18: (A) Schematische Darstellung der Fluoreszenz. Durch Absorption elektromagnetischer Strahlung wird der Fluorophor vom elektronischen Grundzustand (S_0) in den ersten angeregten elektronischen Zustand (S_1) angeregt. Hier befindet er sich zunächst in demjenigen vibratorisch angeregten Zustand, welcher durch maximale Überlappung seiner Wellenfunktion mit der des Ausgangszustandes gekennzeichnet ist (FRANCK-CONDON-Prinzip). Über strahlungslose Übergänge wird Wärmeenergie frei, bis sich der Fluorophor im vibratorischen Grundzustand ($v = 0$) befindet. Von dort aus relaxiert der Fluorophor in den elektronischen Grundzustand (S_0) unter Emission elektromagnetischer Strahlung geringerer Energie. (B) Bekannte Fluorophore mit Angabe der jeweiligen Absorptions/Emissions-Wellenlängen.^[94, 95] Modifizierte Abbildung nach ATKINS et al.^[96]

Selektiv an APs und NFTs bindende Fluoreszenzsonden stellen somit einen vielversprechenden neuen Ansatz für die Alzheimer-Diagnostik *in vivo* über eine Retinauntersuchung dar. Mittels eines sog. konfokalen Laser-Raster-Ophthalmoskops ist es möglich, die Fluoreszenz der Sonden bei verschiedenen Wellenlängen anzuregen, die Emission zu detektieren, und somit ein kostengünstiges und schnelles

in vivo-Imaging zu erreichen.^[97] Es bestehen dabei essentielle Anforderungen an die Fluoreszenzsonden. Sie sollten die Blut-Hirn-Schranke überwinden können, damit sie nicht direkt in den Glaskörper des Auges injiziert werden müssen, nach dem Färben möglichst schnell wieder aus dem Gewebe abtransportiert werden, eine geringe Hintergrundfluoreszenz erzeugen und atoxisch sein. Darüber hinaus sind Quantenausbeute, STOKES-Verschiebung, *Photobleaching*-Verhalten und Löslichkeit für die Eignung eines Fluorophors entscheidend. Damit Fluoreszenzerscheinungen der Sonde nicht überstrahlt werden, sind sog. Anregungsfilter im Einsatz, die langwelliges Licht möglichst vollständig eliminieren. Durch geschickte Wahl des Filters können maximale Signal/Rausch-Verhältnisse erreicht werden. Zu den klassischen Filtern gehört beispielsweise DAPI, dessen Durchlässigkeit auf die Anregung des 4',6-Diamidino-2-phenylindols **10** optimiert wurde (Abbildung 18B). Weitere niedermolekulare Fluorophore sind Fluorescein **11** und 7-Amino-4-methylcoumarin (AMC) **12**. Letzteres kann peptidisch an den C-Terminus spezifischer Enzymsubstrate gebunden werden, um in Protease-Assays eingesetzt zu werden. In gebundenem Zustand wird die Fluoreszenz des AMC **12**-Fragments ausgelöscht und erst nach hydrolytischer Spaltung fluoreszierendes AMC **12** freigesetzt. Dies ist die übliche Methode zur Bestimmung der proteasomalen Restaktivität nach Zugabe eines Inhibitors, um dessen IC₅₀-Wert zu ermitteln. Eine für die AD-Diagnostik etablierte Fluoreszenzsonde ist Methoxy-XO4 **13**. Dieses, von Kongorot und Chrysamin G abgeleitete, biphenylische Distyrylbenzolderivat färbt selektiv NFTs und APs in Hirngewebe.^[98] Nachteile von Methoxy-XO4 **13** sind dessen sehr langsame *Clearence* und schwache Selektivität gegenüber einer der AD-spezifischen Ablagerungen, was eine Optimierung dieser wichtigen Eigenschaften nötig macht.

1.3.4 DYRK1A-Inhibition als potentielle Therapie der AD

Die *dual specificity tyrosine-phosphorylation-regulated kinase 1A* (DYRK1A) gehört zur Gruppe der eukaryotischen CMCG-Kinasen, zu der auch die Glykogen-Synthase-Kinasen 3 (GSK3), die mitogen-aktivierten Proteinkinasen (MAPK), die Cdc2-ähnlichen Kinasen (CLK) und die cyclinabhängigen Kinasen (CDK) gehören.^[99, 100] Die DYRK-Kinasen sind nach ihrer Fähigkeit benannt, direkt nach der Translation eine

Autophosphorylierung an Tyr321 durchzuführen und sich somit selbst in einen aktivierten Zustand zu überführen.

DYRK1A wird mit der Entstehung verschiedener neurodegenerativer Erkrankungen in Verbindung gebracht, darunter die PARKINSON-Krankheit, die HUNTINGTON-Krankheit, die PICK-Krankheit und die ALZHEIMER-Demenz.^[99, 101-103] In einer Studie über den genetischen Einfluss des Chromosoms 21 auf die Pathogenese der AD, an der insgesamt 374 AD-Patienten und eine ebenso große Kontrollgruppe teilnahmen, konnte das DYRK1A-Gen als der genetische Risikofaktor mit der größten Signifikanz identifiziert werden.^[104]

2001 wurde von WOODS et al. erstmals die Bildung von NFTs mit DYRK1A in Zusammenhang gebracht: Die Phosphorylierung von Thr21 des Tau-Proteins war der Auslöser für eine Hyperphosphorylierung durch GSK3.^[105] Zunächst wurden drei Phosphorylierungsstellen an Tau erkannt, die von DYRK1A direkt modifiziert werden: Ser202, Thr212 und Ser404. Mittlerweile wurde nachgewiesen, dass es mindestens elf solcher Serin- und Threoninreste im Tau-Protein gibt.^[86, 106] Durch Phosphorylierung des *alternative splicing factor* (ASF) wird dessen Fähigkeit zur Regulation des Tau-Splittings deaktiviert, was zusätzlich eine Aggregation zu NFTs erleichtert.^[107]

DYRK1A spielt ebenfalls eine Rolle im APP-Metabolismus und somit bei der Bildung amyloider Plaques. Es konnte eine Korrelation zwischen erhöhter DYRK1A-induzierter Phosphorylierung von APP und A β -Leveln beobachtet werden.^[108] Außerdem ist DYRK1A in der Lage die Aktivität der γ -Sekretase durch Phosphorylierung ihrer Untereinheit PS1 zu erhöhen, was ebenfalls zu einer Erhöhung des A β_{42} -Level und somit zu einer vermehrten Bildung von APs führt.^[109] Da DYRK1A sowohl mit der Amyloid-Hypothese (siehe Abschnitt 1.3.1) als auch mit der Tau-Hypothese (siehe Abschnitt 1.3.2) in Zusammenhang steht, gilt es als vielversprechendes neues *Target* zur Behandlung von Morbus ALZHEIMER.

Der bislang potenteste und oral verfügbare DYRK1A-Inhibitor ist Harmin **14** ($IC_{50,DYRK1A} = 80$ nM), dessen Bindungsmodus 2010 von OGAWA et al. mittels Röntgenstrukturanalyse aufgeklärt werden konnte (Abbildung 19A).^[110] Während Harmin **14** eine hohe Selektivität gegenüber anderen Kinasen aufweist, ist es ein potenter Inhibitor der Mono-aminooxidase-A (MAO-A, $K_i = 5$ nM).^[111-113] Diese Ko-Inhibition führt zu der Schlussfolgerung, dass Harmin **14** sowohl psychoaktive

Nebenwirkungen verursacht als auch die Halbwertszeit psychoaktiver Substanzen im Organismus erhöht.

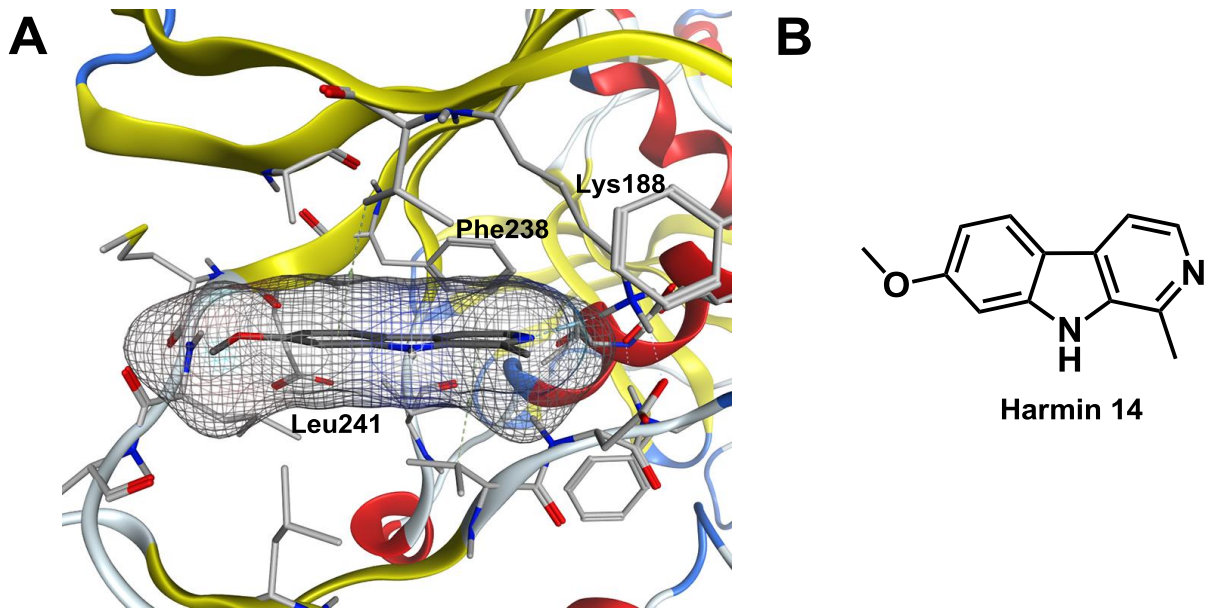


Abbildung 19: (A) Harmin 14 im Komplex mit DYRK1A (PDB-ID: 3anq, visualisiert mit MOE2014.09).^[110] (B) Struktur von Harmin 14.

Sie erklärt außerdem die traditionelle Anwendung eines stark Harmin 14-haltigen Getränks namens *Ayahuasca* in Teilen des Amazonas-Gebietes. Schamanen nutzen die psychoaktiven Effekte des Tranks, um sich in rituelle Trancezustände zu versetzen.^[114] Ein wichtiges Ziel des rationalen Designs neuer DYRK1A-Inhibitoren muss es sein, die Ko-Inhibition von MAO-A zu unterdrücken und dabei die nanomolare inhibitorische Aktivität gegenüber DYRK1A und die orale Verfügbarkeit zu erhalten oder weiter zu verbessern.

2. Zielsetzung

Modernes Wirkstoffdesign beruht auf einem interdisziplinären Zusammenspiel von computerbasierten, chemisch-synthetischen, physikalischen, biochemischen und biologischen Methoden, welches eines zum Ziel hat: Hochaktive und -selektive Verbindungen hervorzubringen, die *in vivo* minimale Nebenwirkungen und optimale pharmakokinetische Eigenschaften zeigen. Gerade diese Interdisziplinarität soll sich in der vorliegenden Arbeit widerspiegeln, weshalb folgende Ziele definiert wurden:

Optimierung von α -Ketophenylamiden als Proteasominhibitoren und mechanistische Studie zur enzymatischen Aktivität bei Ligand/Met45-Kontakt.

In über 30 Jahren Proteasomforschung konnten unzählige Moleküle mit inhibierender Wirkung gegenüber dem 20S-Proteasom identifiziert und synthetisiert werden. Der weitaus größte Anteil fällt hierbei auf die Klasse der kovalenten Inhibitoren, die durch eine elektrophile Kopfgruppe in der Lage sind, eine kovalente Bindung mit dem Thr1 des aktiven Zentrums einzugehen. In einem systematischen Kopfgruppenvergleich konnten STEIN et al. die α -Ketophenylamid-Kopfgruppe als am vielversprechendsten für die Anwendung als Anti-Tumorstoff identifizieren.^[23] Diese Einschätzung beruht auf der äußerst hohen inhibitorischen Aktivität gegenüber der β 5-Untereinheit bei gleichzeitigem reversiblen Bindungsmodus. Die klinisch etablierten Inhibitoren Bortezomib und Carfilzomib sind hingegen (pseudo)-irreversible Binder, was eine Penetration und Behandlung von solidem Tumorgewebe ausschließt. In einem iterativen Design/Synthese-Prozess soll mit Hilfe von *in silico*-Methoden, wie beispielsweise Docking, QSAR und Simulation der Moleküldynamik, die Leitstruktur BSc2189 9 hinsichtlich ihrer Aktivität gegenüber der β 5-Untereinheit optimiert werden. Im Anschluss an eine erfolgreiche Optimierung ist die Charakterisierung des Bindungsmodus des neuen Wirkstoffes anhand röntgenkristallographischer Untersuchungen und Dialyseexperimenten geplant. Weiterhin ist eine mechanistische Studie Ziel dieser Arbeit, die Aufschluss über den Einfluss eines Kontaktes des Liganden mit der Aminosäure Met45 auf die enzymatische Aktivität der β 5-Untereinheit geben soll.

Entwicklung einer Methode zum automatisierten Docking kovalenter Liganden

Eine der meistgenutzten computerbasierten Methoden im Rahmen des Wirkstoffdesigns stellt zweifellos das Protein/Ligand-Docking dar. Softwareentwickler setzten ihren Fokus in der Vergangenheit nahezu ausschließlich auf die Simulation nicht-kovalenter Liganden. Dies hat zur Folge, dass die wenigen Programme, die mit einer kovalenten Bindung zwischen Ligand und Rezeptor umgehen können, hinsichtlich ihrer Benutzerfreundlichkeit, Rechenintensität, Automatisierung und Vielseitigkeit erhöhten Optimierungsbedarf erkennen lassen. Das Softwarepaket *Molecular Operating Environment* (MOE, Chemical Computing Group) stellt eine umfangreiche Softwareumgebung für das Wirkstoff- und Proteindesign dar und ermöglicht darüber hinaus die Implementierung eigener Programme mit Hilfe der MOE-exklusiven Programmiersprache *Scientific Vector Language* (SVL).^[60, 115] Im Rahmen dieser Arbeit soll eine Methode zum kovalenten Docking entwickelt werden, die keinerlei Restriktionen bezüglich elektrophiler Kopfgruppen oder Rezeptorklassen zeigt, benutzerfreundlich und dennoch bei Bedarf leicht modifizierbar ist. Ferner soll die Methode mit möglichst wenig Rechenleistung auskommen, automatisiert ablaufen und somit virtuelle Screenings größerer Liganden-Datenbanken ermöglichen.

Entwicklung einer optimierten Syntheseroute von Methoxy-X04-basierten Fluoreszenzfarbstoffen zur Diagnose der ALZHEIMER-Demenz

Morbus ALZHEIMER wird durch zwei Proteinablagerungen im Gehirn charakterisiert: Die extrazellulären amyloiden Plaques (APs) und die intrazellulären Neurofibrillenbündel (NFTs), die auf eine hyperphosphorylierte Form des Tau-Proteins zurückzuführen sind. Bildgebende, nicht-invasive Verfahren zur Identifikation dieser Ablagerungen sind vielversprechende Hilfsmittel, um eine akute oder bevorstehende Erkrankung an Morbus Alzheimer sicher diagnostizieren und letztlich Therapiemaßnahmen einleiten zu können. Neben den apparativ aufwendigen Verfahren, wie der Positronen Emissions- (PET), der Einzelphotonenemissions- (SPECT) oder der Computer-Tomographie (CT), ist vor allem die Fluoreszenzmikroskopie ein wichtiger Forschungsschwerpunkt, mit deren Hilfe pathogene Ablagerungen, beispielsweise in der Retina, zukünftig im Rahmen von Vorsorgeuntersuchungen nachgewiesen werden sollen. Da das Vorhandensein von APs, im Gegensatz zu NFTs, nicht zwingend mit kognitiven Störungen durch eine

bestehende ALZHEIMER-Demenz einhergeht, besteht ein großes Interesse an NFT-selektiven Fluoreszenzsonden. Ausgangspunkt für eine Derivatisierungsstudie stellt der etablierte Fluoreszenzfarbstoff Methoxy-X04 dar. Im Rahmen dieser Arbeit soll eine im Hinblick auf Ausbeute und Syntheseaufwand optimierte Syntheseroute entwickelt und mit ihrer Hilfe eine Vielzahl neuer fluoreszenter Methoxy-X04-Derivate mit potentiell erhöhter Selektivität gegenüber einer der genannten pathogenen Ablagerungen hergestellt werden.

Untersuchung der Struktur/Aktivitäts-Beziehung (SAR) von β -Carbolinen als DYRK1A-Inhibitoren und Steigerung der Selektivität gegenüber MAO-A mit Hilfe moderner *in silico*-Methoden.

Die Kinase *dual specificity tyrosine-phosphorylation-regulated kinase 1A* (DYRK1A) konnte sowohl mit der Amyloid- als auch mit der Tau-Hypothese in Verbindung gebracht werden und spielt somit in beiden Erklärungsansätzen der Pathogenese von Morbus ALZHEIMER eine wichtige Rolle. Mit Hilfe von computerbasierten Methoden, wie Pharmakophor-Modellierung und Protein/Ligand-Docking, soll die Struktur/Aktivitäts-Beziehung (SAR) von β -Carbolinen als DYRK1A-Inhibitoren aufgeklärt und durch ein Zusammenspiel von chemischer Synthese und biologischer Evaluation geeigneter Derivate validiert werden. Übergeordnetes Ziel ist es, basierend auf den Erkenntnissen der SAR-Untersuchung, β -Carbolinderivate mit erhöhter Selektivität gegenüber der Mono-aminooxidase-A (MAO-A) zu synthetisieren. Die Ko-Inhibition von MAO-A ist ursächlich für eine Reihe von erheblichen Nebenwirkungen des bislang aktivsten β -Carbolin-basierten DYRK1A-Inhibitors Harmin **14**.

3. Kumulativer Teil

3.1 Optimierung von α -Ketophenylamiden als Proteasominhibitoren

Der Inhalt dieses Kapitels wurde bereits veröffentlicht:

Constantin Voss, Christoph Scholz, Sabine Knorr, Philipp Beck, Martin L. Stein, Andrea Zall, Ulrike Kuckelkorn, Peter M. Kloetzel, Michael Groll, Kay Hamacher, Boris Schmidt, "alpha-Keto Phenylamides as P1'-Extended Proteasome Inhibitors", *ChemMedChem* **2014**, 9, 2557-2564.

Mit freundlicher Genehmigung des Verlags Wiley-VCH. Nummerierungen wie in der Publikation.

Wie in Abschnitt 1.1.4 beschrieben, bieten peptidische α -Ketophenylamide aufgrund ihres reversiblen Bindungsmodus die Möglichkeit, tiefliegendes Tumorgewebe zu penetrieren. Die Ausweitung der Proteasominhibition als Therapiestrategie auf solide Tumoren könnte ausschlaggebend für den klinischen Erfolg von Wirkstoffen der nächsten Generation sein. α -Ketophenylamide verfügen durch den Kontakt mit der gestrichenen Seite des Substratbindungskanals zudem über zusätzliche Möglichkeiten, Wechselwirkungen mit dem aktiven Zentrum einzugehen und die Selektivität gegenüber dem Proteasom zu erhöhen.

Ausgangspunkt der vorliegenden Wirkstoffoptimierung ist die Leitstruktur BSc2189 **5**, welche über einen unsubstituierten Phenylamidterminus verfügt und bereits eine beachtliche Inhibition der ChTL-Aktivität mit einem IC₅₀-Wert von 72 nM zeigt. Aus der röntgenkristallographischen Analyse von BSc2189 **5** im Komplex mit dem Proteasom war bekannt, dass der Phenylring eine nahezu perfekt koplanare Ausrichtung bezüglich der Amidbindung einnimmt. Aus dieser Beobachtung folgte die Arbeitshypothese, dass die gehinderte Rotation des Phenylrings auf eine Konjugation des aromatischen Systems mit der Amidbindung zurückzuführen ist, die bekanntermaßen erhöhten sp²-Charakter besitzt. In diesem Fall wäre die Rotation schon vor Bindung an das aktive Zentrum gehindert, was zu einer entropisch begünstigten Ligandenbindung führt. Um diese Hypothese zu belegen, wurde eine quantitative Struktur-Aktivitäts-Beziehung (QSAR) aufgestellt, welche die HAMMETT-Konstante σ als elektronischen Parameter für die Elektronendichte am *ipso*-Kohlenstoffatom des Phenylterminus beinhaltet (siehe Abschnitt 1.2.2). Hierzu wurden insgesamt vier Derivate von BSc2189 **5** synthetisiert: Je zwei mit

Elektronendonorguppe (**6b** und **6c**) und -akzeptorguppe (**6d** und **6e**) in *para*-Position. Die Auftragung der experimentell ermittelten apparenten Bindungsaffinitäten (pIC_{50}) gegen die tabellierten HAMMET-Konstanten σ zeigte den erwarteten linearen Trend, dass eine erhöhte Elektronendichte am *ipso*-Kohlenstoffatom zu einer erhöhten inhibitorischen Aktivität gegenüber dem Proteasom führt. Diese Beobachtung kann mit einem erhöhten Doppelbindungscharakter und einer damit einhergehenden erhöhten Rigidität im ungebundenen Zustand begründet werden.

Eine Optimierung der inhibitorischen Aktivität der Leitstruktur konnte jedoch mit keiner der durchgeführten Derivatisierungen erreicht werden. Um dies zu erklären, wurden neben den intramolekularen nun auch intermolekulare Wechselwirkungen in die Arbeitshypothese aufgenommen. Alle eingefügten *para*-Substituenten waren polarer Natur und in der Lage ein H-Brückennetzwerk mit vorher freiem Wasser im inneren Zylinder des 20S-Proteasoms zu bilden, was den entropischen Vorteil der Ligandenrigidität überkompensieren würde. Der Fokus richtete sich daraufhin auf Methylsubstituenten, die als σ -Donorgruppen zwar die Elektronendichte des aromatischen Systems erhöhen, jedoch keine H-Brücken eingehen können. Den größten Einfluss auf die Elektronendichte am *ipso*-Kohlenstoff besitzen *ortho*- und *para*-Substituenten, von denen es insgesamt fünf verschiedene Substitutionsmuster gibt (2, 4, 2/4, 2/6 und 2/4/6). Um den chemisch-synthetischen Aufwand so gering wie möglich zu halten, wurde ein kovalentes Docking der fünf Derivate in das aktive Zentrum der $\beta 5$ -Untereinheit durchgeführt. Als vielversprechendstes Derivat wurde BSc4999 **6a** (2,4-Dimethyl) identifiziert und daraufhin synthetisiert.

Mit Hilfe dieses interdisziplinären Ansatzes aus röntgenkristallographischer Untersuchung des Bindungsmodus, Aufstellen einer QSAR-Gleichung, kovalentem Protein/Ligand-Docking und einer detaillierten Evaluation mittels biochemischer Assaymethoden, konnte mit BSc4999 **6a** ($IC_{50} = 38 \text{ nM}$) eine signifikante Steigerung der inhibitorischen Aktivität mit minimalem Synthesaufwand erreicht werden. Hierbei wurden der reversible Bindungsmodus und die hohe $\beta 5$ -Selektivität der Leitstruktur beibehalten und mit Einführung zweier Methylgruppen in der S1'-Position neue Selektivitätskriterien zur möglichen Reduktion unerwünschter Nebenwirkungen, wie der Peripheren Neuropathie, eingefügt.

α -Keto Phenylamides as P1'-Extended Proteasome Inhibitors

Constantin Voss,^[a] Christoph Scholz,^[a] Sabine Knorr,^[b] Philipp Beck,^[c] Martin L. Stein,^[c] Andrea Zall,^[a] Ulrike Kuckelkorn,^[d] Peter-Michael Kloetzel,^[d] Michael Groll,^[c] Kay Hamacher,^[b] and Boris Schmidt^{*[a]}

The major challenge for proteasome inhibitor design lies in achieving high selectivity for, and activity against, the target, which requires specific interactions with the active site. Novel ligands aim to overcome off-target-related side effects such as peripheral neuropathy, which is frequently observed in cancer patients treated with the FDA-approved proteasome inhibitors bortezomib (**1**) or carfilzomib (**2**). A systematic comparison of electrophilic headgroups recently identified the class of α -keto amides as promising for next generation drug development. On the basis of crystallographic knowledge, we were able to develop a structure–activity relationship (SAR)-based approach for rational ligand design using an electronic parameter (Ham-

mett's σ) and in silico molecular modeling. This resulted in the tripeptidic α -keto phenylamide BSc4999 [(S)-3-(benzyloxycarbonyl-(S)-leucyl-(S)-leucylamino)-5-methyl-2-oxo-N-(2,4-dimethylphenyl)hexanamide, **6a**], a highly potent (IC_{50} = 38 nM), cell-permeable, and slowly reversible covalent inhibitor which targets both the primed and non-primed sites of the proteasome's substrate binding channel as a special criterion for selectivity. The improved inhibition potency and selectivity of this new α -keto phenylamide makes it a promising candidate for targeting a wider range of tumor subtypes than commercially available proteasome inhibitors and presents a new candidate for future studies.

Introduction

Intracellular protein degradation via the ubiquitin–proteasome system is crucial for protein homeostasis within eukaryotic organisms. The 26S proteasome is a sophisticated multicatalytic molecular degradation machinery consisting of a proteolytic 20S core particle (CP) surrounded by two regulatory 19S caps, which are responsible for the recognition of ubiquitin marked substrates, their unfolding, and transport to the inner CP. The 20S proteasome appears as an elongated hollow cylinder forming the multicatalytic center: four heptameric rings form a pile following an $\alpha_2\beta_2\beta_2\alpha_2$ stoichiometry.^[1] The binding channels of the proteasome active sites within the inner β -rings bear different subpockets that are essential for the specific binding of defined substrate peptides. A systematic nomenclature differenti-

ates the primed and a non-primed regions of this channel, starting from the cleavage site of a peptide substrate.^[2]

However, fluorogenic assays with eukaryotic cells revealed that each β -ring only harbors three catalytic active sites with distinct substrate preferences: $\beta 1$ cleaves after acidic residues and therefore mimics a caspase-like (CL) activity, $\beta 2$ displays trypsin-like (TL) activity, while $\beta 5$ typifies chymotrypsin-like (ChTL) activity.

The mode of action of each active site follows a uniform mechanism employing an N-terminal threonine (Thr1) hydrolyzing the substrate's scissile peptide bond by nucleophilic attack of its hydroxy group (Thr1O⁻). The N terminus (Thr1N) acts via a water molecule and coordinates the proton shuttle and cleavage of the acyl–enzyme intermediate by the release of defined oligopeptides with a length distribution between three and 25 amino acids.^[3] Due to the central role of the CP in antigen processing, cell cycle control, cell signaling, and protein quality control, this protease represents an important target in the fields of cell biology, structural biology, and medicinal chemistry. Numerous specific and nonspecific inhibitors have been developed to target proteasome activities.^[4] This led to the approval of the boronic acid bortezomib (**1**) by the US Food and Drug Administration (FDA) in 2003 and, most recently, the epoxyketone carfilzomib (**2**) for the treatment of multiple myeloma in 2012 (Figure 1).^[5]

So far, the majority of known natural or synthetic inhibitors of the proteasome addresses the non-primed site of the binding channel. However, the primed regions show significant variations amongst the active sites and thus can be used as selec-

[a] C. Voss, C. Scholz, Dr. A. Zall, Prof. Dr. B. Schmidt
Clemens Schöpf Institute for Organic Chemistry & Biochemistry
Technische Universität Darmstadt
Alarich Weiss-Str. 4-8, 64287 Darmstadt (Germany)
E-mail: schmidt_boris@t-online.de

[b] S. Knorr, Prof. Dr. K. Hamacher
Computational Biology & Simulation, Technische Universität Darmstadt
Schnittspahnstr. 10, 64287 Darmstadt (Germany)

[c] P. Beck, Dr. M. L. Stein, Prof. Dr. M. Groll
Center of Integrated Protein Science at the Department of Chemistry
Chair of Biochemistry, Technische Universität München
Lichtenbergstr. 4, 85747 Garching (Germany)

[d] Dr. U. Kuckelkorn, Prof. Dr. P.-M. Kloetzel
Institute of Biochemistry CCM, Charité Universitätsmedizin Berlin
Charitéplatz 1/Virchowweg 6, 10117 Berlin (Germany)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cmdc.201402244>.

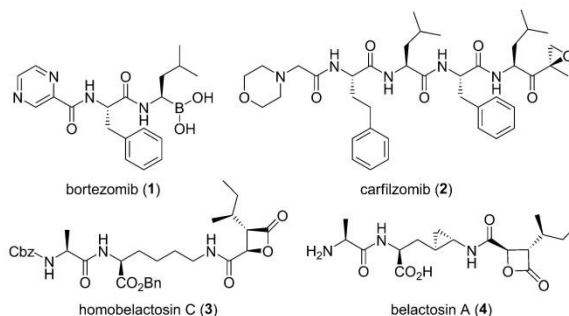


Figure 1. Structures of FDA-approved proteasome inhibitors **1** and **2**, as well as **3**, which binds to the primed site of the proteasome substrate binding channel, and its natural product counterpart **4**.

tivity criteria for structure-guided inhibitor design. As an example, the β -lactone proteasome inhibitor homobelactosin C (**3**, Figure 1) shows high β 5-selectivity by occupying the primed site of the ChTL substrate binding channel. Interestingly, this natural product displays high antitumor activity, with nanomolar IC_{50} values in human pancreoma and colon cancer cells.^[6] Crystallographic analysis of this ligand in complex with the yeast 20S proteasome identified that its high selectivity was caused by an unexpected mode of action.^[7] Thus, **3** and its natural counterpart, belactosin A (**4**, Figure 1), turned out to be promising leads for the design of new and highly selective proteasome inhibitors.^[8] Even though acyl-ester formation of the β -lactone headgroup with Thr10^Y follows a reversible mechanism, the ligand orientation blocks access to a water molecule, preventing deacetylation. The resulting long half-life of β -lactone-induced inhibition is similar to the activity of irreversibly binding ligands. This imposes limits for deep solid tissue penetration, restricting these inhibitors to the treatment of non-solid tumors.^[9]

Recently, the class of reversible and potent peptidic α -keto amides has been shown to exploit both the primed and the non-primed sites, thereby compensating for their moderate chemical reactivity and leading to a strong proteasome binding preference.^[10] It turned out that the α -keto phenylamide headgroup harbors a potential motif for ligand optimization, due to its unique orientation in the CP active site cavity with the terminal phenylamide moiety projecting into the primed site of the binding channel. Furthermore, it is the α -keto phenylamide residue that accounts for the ligand's inhibitory potency.

Therefore α -keto phenylamide CP inhibitors are promising candidates for preclinical studies of a wider range of tumor subtypes as currently targeted by bortezomib (**1**) and carfilzomib (**2**). This work describes the development of an inhibitor based on the α -keto phenylamide headgroup, that engages in additional primed site interactions and thus exhibits improved potency.

Results

In a previous work,^[10a] BSc2189 (**5**, Figure 2A) was identified as a promising potent proteasome inhibitor (IC_{50} : 72 nM) that exhibits high selectivity for the β 5 subunit. Diminished proteolysis in the cytosolic fraction, as well as in the protease inhibitor mixture-pretreated lysate (Complete, Roche Applied Science), confirmed inhibitory activity exclusively to the CP and distinguished it from inhibition of most cytosolic serine/aspartate proteases. X-ray data from crystallization of the compound in complex

with the yeast 20S proteasome revealed hemiacetal formation in the α -position of the phenylamide moiety, due to nucleophilic attack of the ligand α -ketone by Thr10^Y (Figure 2A).^[10c] The observation of the amide terminus occupying the S1' cavity of the binding channel is in accordance with a previous work from Chatterjee et al.^[10b] X-ray data revealed the ligand's primed site residue (P1') extending into the hollow inner side of the cylindrical CP. The attack by Thr10^Y on the ligand's "si" face enables hydrogen bond formation between the hemiacetal and Thr1N (Figure 2B). Occupation of the oxyanion hole formed by Gly47N with the terminal amide carbonyl group additionally stabilizes the ligands orientation on the target. However, the most notable feature of **5** is its aromatic C-terminal phenylamide moiety that resides in an almost perfectly planar fashion in the S1' subpocket of the binding channel (Figure 2D). This rigidity of the P1' residue of the ligand contributes to the driving force for its high-affinity binding, as decreased degrees of freedom in the unbound state restrict the entropic penalty upon binding.

As a proof of concept, we investigated the variation in the electron density of the aromatic system by introducing substituents at the *para* position of the phenyl moiety in P1'. Both electron donating (**6b**; **6c**) and electron withdrawing groups (**6d**; **6e**) were analyzed in this study. We converted the respective anilines (**7a–d**) to formamides (**8a–d**, Scheme 1) to provide access to functionalized derivatives of **5**. N-formylation was performed according to a published catalyst- and solvent-free procedure, which was adopted for anilines **7a–d**.^[11] The subsequent conversion into isonitriles **9a–e** was followed by a dehydration step, using phosphorus oxychloride and triethylamine in dichloromethane.^[12] Subsequent conversion was done directly after a quick workup of hydrolytically labile compounds **9a–e**. The tripeptidic aldehyde Cbz-Leu-Leu-Leu-al and trifluoroacetic acid then underwent a multicomponent Passerini reaction with isonitriles **9a–e** to form α -hydroxyphenylamides **10a–e**. Oxidation with 2-iodoxybenzoic acid (IBX) provided the desired tripeptidic α -keto amides **6a–e** in moderate yields, as displayed in Scheme 1.

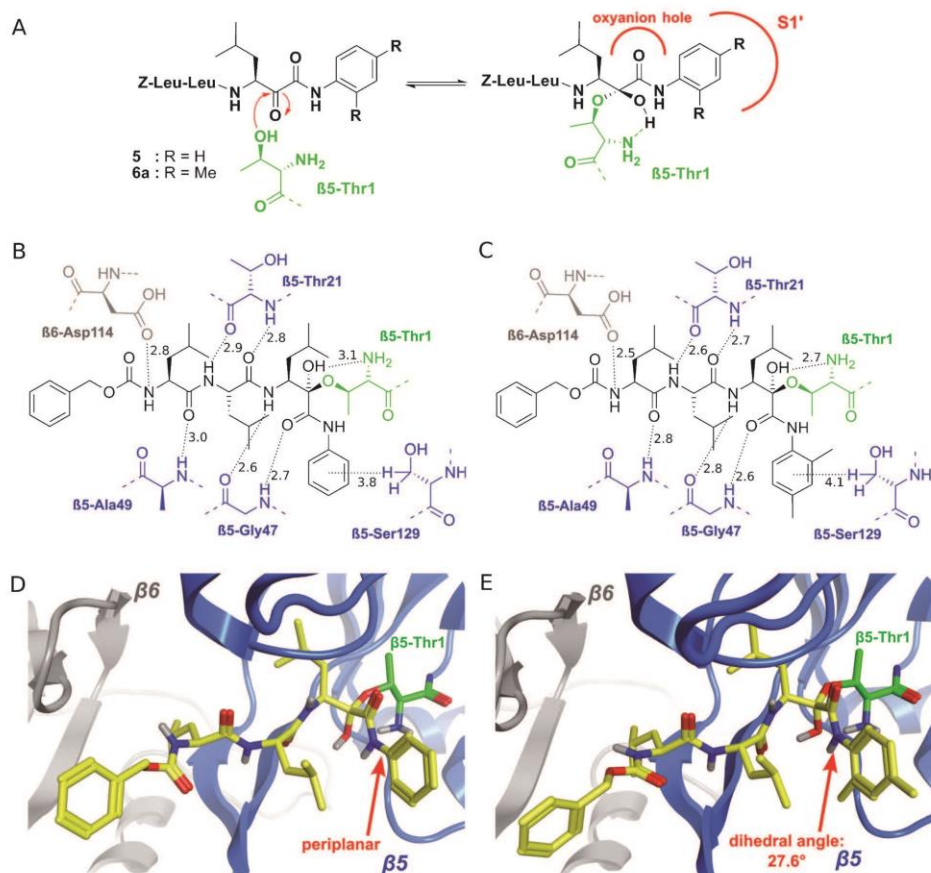
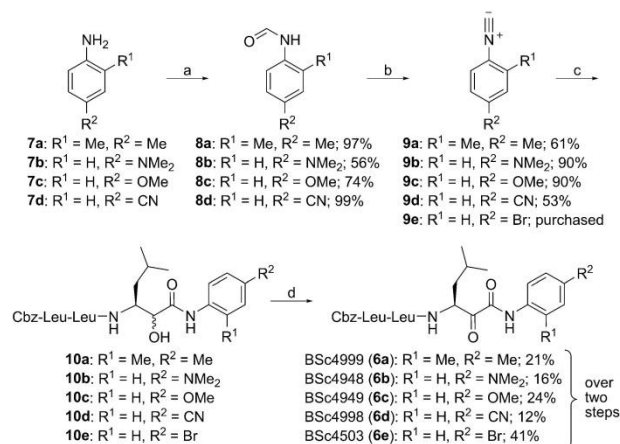


Figure 2. A) Binding mechanism of α -keto phenylamides to the $\beta 5$ subunit of the 20S proteasome. Reversible hemiacetal formation by nucleophilic attack of the catalytically active Thr 1 (green) with the ketone moiety of the inhibitor (black). B)–E) Crystal structures of B5c2189 (5, PDB code: 4N08) and B5c4999 (6a, PDB code: 4R02) in complex with the yeast CP. B, C) 2D images show interactions as grey dashed lines between the inhibitor and relevant subpocket residues of the proteasome (blue/grey). D, E) 3D images show the ligands as stick models (yellow) and the proteasome subunits in ribbon/loop representation (blue/grey). The dihedral angles ($C_{\text{carbonyl}}-N-C_{\text{phenyl}}-C_{\text{phenyl}}$) of the ligand phenyl amide moieties are indicated by red arrows.

A structure–activity relationship study was established for our ensemble of initial derivatives **6b–e**, using the tabulated Hammett's constants (σ), a parameter expressing the electronic constitution of an aromatic reaction center in dependency of the nature of a substituent at a *meta* or *para* position.^[13] Our hypothesis is that the binding process to the CP predominantly depends on the electron density of the aromatic C-termini, modified by different *para* substituents in **6b–e**. Therefore, we performed a hypothesis test in which we plotted logarithmic IC_{50} values against the respective σ constants. As a result, we identified a linear correlation supporting our hypothesis: biological data were consistent with the SAR for all *para*-substituted compounds (**6b–e**), demonstrating a loss in activity with higher σ constants (Figure 3).

Accordingly, we observed that increased electron density of the aromatic system, which is induced by electron donating substituents, resulted in a stronger double bond character of the nitrogen–phenyl bond ($N-C_{\text{phenyl}}$) and, thus, to increased rigidity. We assume this to be the major feature of an entropically favored binding process, thus corresponding to the improved biological activity. Nevertheless, all derivatives carrying functional groups in the *para* position were less active than the unsubstituted phenylamide lead (**5**). For illustration purposes, we have included **5** in Figure 3. Hence, we aimed to characterize the intermolecular forces within the primed site of the $\beta 5$ subunit. The SAR study suggested that the electron donating groups play a key role in the affinity of the inhibitor, contrary to the withdrawing groups. As the considered elec-



Scheme 1. Synthesis of target compounds **6a–e**. Reagents and conditions: a) HCO₂H, 60 °C; b) POCl₃, NEt₃, CH₂Cl₂, 0 °C; c) 1. Cbz-Leu-Leu-OH, pyridine, CH₂Cl₂, –10 °C, 2. TFA, 0 °C; d) IBX, DMSO, RT. Percent yields are shown; for **6a–e**, yields were calculated over two reaction steps, as **10a–e** were not isolated.

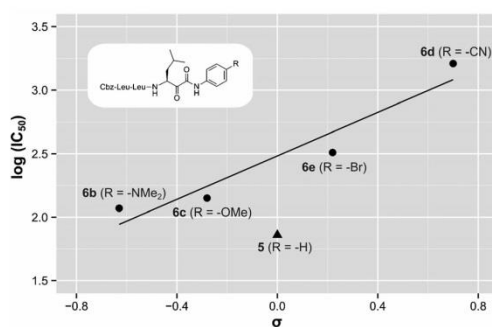


Figure 3. Plot of logarithmic IC₅₀ values of inhibitors **6b–e** against respective Hammett's constants σ (●). Note the significant linear correlation ($p \leq 0.0409$) between $\log(\text{IC}_{50})$ and σ . For illustrative purposes, we also added a single data point for **5** (▲). Significance tests were performed based on our starting hypothesis for exclusively *para*-substituted derivatives **6b–e**. Linear correlation: $\log(\text{IC}_{50}) = 0.85584\sigma + 2.48202$.

tron donating groups of **6b** and **6c** predominantly act as hydrogen bond acceptors, water molecules are coordinated by each of the ligands and cause a decrease in their binding preference by compensating for the benefit of high electron density within the aromatic system. We therefore propose that methyl groups donate electron density into the aromatic system by α -conjugation, but they do not participate in hydrogen bonding. We performed a docking study to evaluate the most promising compound of the five possible *ortho*- and *para*-substituted methylphenylamide derivatives. Covalent docking with flexible side chains of the receptor was realized with *conflexdock* within the MOE2012.10 software.^[14] First,

a conformational database spanning different dihedral angles between the plane of the respective aromatic system and the amide moiety was generated by keeping the peptidic backbone in its “native conformation”, based on the blueprint of the CP-5 X-ray structure. Next, simulations starting with iterated dihedral angles in 45° steps were carried out to determine relative energy minima, allowing us to rank the ligands according to their respective energies (see Supporting Information). The 2,4-dimethyl substitution displayed the lowest overall energy score, suggesting the synthesis of ligand **6a**. In agreement with our modeling studies, **6a** turned out to be the most potent inhibitor of the 20S proteasome within this series. Blocking the

ChTL activity *in vitro* resulted in an IC₅₀ value of 38 nM. Moreover, this compound is selective for the ChTL substrate binding channel, as the CL and the TL activities remained unaffected (Table 1).

Table 1. Activities of peptidic α -keto phenylamides **6a–e** against the three catalytic activities of isolated proteasomes.^[a]

Compd	ChTL ($\beta 5$)	IC ₅₀ [nM] ^[b]	
		TL ($\beta 2$)	CL ($\beta 1$)
6a	38 ± 19	> 3000	NA
6b	118 ± 40	NA	NA
6c	142 ± 95	> 9000	NA
6d	1608 ± 53	NA	NA
6e	326 ± 35	NA	NA

[a] See Scheme 1 for compound structures. [b] Data, normalized to controls, are the means of two independent experiments, each performed in triplicate ($n = 2$); *in vitro* data for lead compound **5** (IC_{50, $\beta 5$} = 72 nM) were published previously.^[10a] NA: not affected (IC₅₀ > 10 μ M).

Crystallographic data of **6a** in complex with the yeast 20S proteasome revealed that the dimethylated aromatic system is twisted out of the amide plane, with a dihedral angle of 27.6°, in contrast with the co-planar keto amide moiety of lead **5** upon CP binding (Figure 2D,E). This conformational change is interpreted as an on-target effect due to steric and enthalpic interactions within the active site, which compensates for the electronically favored periplanar shape. Note, these findings are in agreement with our predictions, as a hindered rotation of the C_{carbonyl}–N–C_{phenyl}–C_{phenyl} plane prior to bond formation with Thr107 is crucial for the entropically favored binding of the ligand.

Finally, we performed a HeLa cell-based assay to investigate cell permeability of the synthesized compounds *in vivo* (Supporting Information). We obtained promising inhibition activities, thus confirming cell penetration of **5** and **6a–e**. In addition, the reversibility of the α -keto phenylamide binding mode was addressed by a dialysis experiment with **5** and the most potent inhibitor, **6a** (Supporting Information). After inhibitor treatment, the proteasome activity recurred significantly after 72 h, as expected for hemiacetal formation of the ligand with the active sites.

Discussion

Chatterjee *et al.*^[10b] first reported α -keto amides as P1'-extended proteasome inhibitors in 1999, yet crystallographic evaluation of the binding mode to the active site of the CP enabled detailed SAR studies only recently.^[10c] Focusing on the electronic situation of the ligand's aromatic C-terminus and its effect on binding to the CP, we plotted tabulated Hammett's constants from aromatic *para* substituents against IC₅₀ values of the synthesized α -keto phenylamides **6b–e**. We found a correlation between selected substituents of inhibitors and their biological activities, which depend on the strength of the electron donating group. Though the lead structure (**5**) was still the most active inhibitor, compared with **6b–e**, it is the polar nature of the *para*-NMe₂ or *para*-OMe groups that diminishes the biological activity. However, in contrast to nonpolar ligand interactions with lipophilic protein surfaces, the influence of hydrogen bond formation in the thermodynamic ligand binding process is more complex and still needs further experimental characterization.^[15] We thus speculate that a polar group in P1' may induce water binding and create a hydrogen bond network that entropically hinders the ligand binding process rather than stabilizes the on-target structure. By introducing methyl groups, which are nonpolar α -donors, we were able to confirm our conclusions from the SAR studies. Subsequently, an optimal substitution pattern of methyl groups was identified by a molecular docking approach, resulting in the 2,4-dimethylated derivative **6a**. With **6a** having an IC₅₀ value of 38 nM, we could demonstrate a significant enhancement of ligand interaction and inhibitory activity relative to **5** (IC₅₀ = 72 nM). Remarkably, analysis of the crystal structure of **6a** in complex with the yeast proteasome revealed a dihedral twist of the planar conjugated π -system, which is due to increased hydrophobic interactions with the P1' binding site. This deviation from planarity stands in contrast with the rigid and planar arrangement of **5** in complex with the CP. Addressing the primed and non-primed regions of the binding channel differentiates α -keto phenylamides from the majority of known peptidic proteasome inhibitors, thus providing access to highly potent, reversible, and specific small molecule inhibitors. Furthermore, our data enable targeted ligand design in both directions of the substrate binding channel, using the α -keto phenylamide moiety as a superb linker between the primed and non-primed sites.

Conclusions

Identification and biological evaluation of the tripeptidic α -keto phenylamide **5** was the major subject of prior work, including binding mode elucidation, IC₅₀ and LD₅₀ determination, and selectivity profiling of the distinct proteasome active sites and most cytosolic serine/aspartate proteases.^[10a,c] The comparative study of different headgroup inhibitors, all bearing the Cbz-Leu-Leu-Leu-backbone, indicated ketoamide **5** as the most promising candidate to be investigated in cellular models of chemo- and immunosuppressive therapies. The aim of this study was to improve both potency and ligand efficiency of **5** while avoiding significant enlargement, as this may result in too lipophilic and thus undruggable compounds. Our SAR approach identified structure **6a** and provided improved ligand efficiency by marginal modifications of lead structure **5**. Biological evaluation of **6a** showed that the beneficial properties of the α -keto phenyl amide moiety could be conserved, while gaining improved potency against the $\beta 5$ subunit. The S1' occupation of the ligand as an additional selectivity criterion may diminish off-target side effects in humans, such as peripheral neuropathy. This disabling neuropathy is frequently observed after bortezomib treatment and even with the second generation drug carfilzomib. Furthermore, the reversible binding mode of hemiketal formation is likely to enable penetration of deeper solid tissues and allow cells to recuperate unless they are sufficiently damaged. The CP inhibitor **6a** is a promising drug candidate to address a wider range of tumor subtypes than targeted by irreversible commercial drugs, owing to its strong inhibition potency (IC₅₀ = 38 nM), and thus has emerged as a target for further investigations.

Experimental Section

Synthetic procedures

This section contains the experimental description of new compounds synthesized in this work: **6a–e**. Experimental descriptions of synthesized substrate compounds already similarly described in the literature (**8a–d**, **9a–d**) can be found in the Supporting Information. Compound **5** was synthesized by a previously published procedure.^[10a] The aldehyde substrate Cbz-Leu-Leu-Leucinal was synthesized by standardized peptide coupling from commercial enantiopure Cbz-Leu-Leu-OH and (S)-leucinol. Oxidation with IBX gave the desired aldehyde. IBX was synthesized following a published procedure.^[16] All chemicals that were purchased as reagent grade from commercial suppliers were used without further purification.

General methods: ¹H NMR and ¹³C NMR spectra were recorded with a Bruker AC 300 (300 MHz) or AC 500 (500 MHz) spectrometer. Chemical shifts are reported in δ (ppm), adjusted to the central line of the deuterated solvent (MeOD, CDCl₃, D₂O, DMSO). High resolution mass spectrometry (HRMS) was performed with an Agilent 1290 Infinity HPLC system coupled to an Agilent G6530A QTOF MS system. HPLC analysis was performed with an Agilent 1100 system. The purity of the final compounds was determined by UV detection (λ = 254 nm). The chromatographic method employed the following: Zorbax Eclipse XDB-C18 column, 4.6 \times 150 mm; mobile phase A: H₂O (0.1% TFA), mobile phase B:

acetonitrile; flow rate: 1 mL min⁻¹; gradient elution: 30 to 100% B over 15 min. According to this method, the purities for all compounds that were evaluated in biological assays were $\geq 95\%$. Thin-layer chromatography was carried out using aluminum sheets pre-coated with silica gel 60 F254 (0.2 mm; E. Merck). Chromatographic spots were visualized by UV and/or by spraying with a methanolic solution of vanillin/H₂SO₄ or aqueous KMnO₄ solution, followed by heating. Silica gel chromatography was carried out using Merck silica gel 60 (0.063–0.2 mm).

α -Keto phenylamides (6a–e) from peptidic aldehydes and phenyl isonitriles via a Passerini reaction and subsequent oxidation of intermediate α -hydroxy phenylamides (10a–e): The peptidic aldehyde (1.0 equiv), phenyl isonitrile (1.5 equiv), and pyridine (4.0 equiv) were dissolved in dry CH₂Cl₂ (2 mL mmol⁻¹ aldehyde) and cooled to -10°C . Trifluoroacetic acid (2.0 equiv) was added dropwise, and the reaction mixture was allowed to stir for 2 h at 0°C . After stirring for an additional 72 h at room temperature, completion of the reaction was monitored by HPLC. CH₂Cl₂ was added, and the mixture was washed with 0.1 N aqueous HCl (3 \times) and aqueous saturated NaHCO₃ (3 \times). The organic layer was then dried over Na₂SO₄, filtered, and concentrated under reduced pressure. The resulting colorless oil and IBX (1.5 equiv) were dissolved in DMSO (2 mL mmol⁻¹ aldehyde) and stirred for 12 h at room temperature. After addition of CH₂Cl₂, the mixture was washed with water (3 \times), aqueous saturated NaHCO₃ (3 \times), and brine (3 \times). The organic layer was then dried over Na₂SO₄, filtered, and concentrated under reduced pressure. Purification was done via liquid chromatography.

(S)-3-(Benzyloxycarbonyl-(S)-leucyl-(S)-leucylamino)-5-methyl-2-oxo-N-(2,4-dimethylphenyl)hexanamide (6a): Yield: 21% (35 mg), colorless oil: ¹H NMR (300 MHz, CDCl₃, 300 K): δ = 8.57 (1 H, s), 7.85 (1 H, d, J = 8.7 Hz), 7.32 (5 H, m), 7.05 (1 H, d, J = 7.4 Hz), 7.01 (2 H, m), 6.81 (1 H, m), 5.49 (1 H, m), 5.40 (1 H, m), 5.08 (2 H, m), 4.54 (1 H, m), 4.24 (1 H, m), 2.28 (3 H, s), 2.24 (3 H, s), 1.58 (9 H, m), 0.92 ppm (18 H, m); ¹³C NMR (125 MHz, CDCl₃, 300 K): δ = 196.9, 172.6, 171.9, 156.8, 156.3, 136.2, 135.6, 131.7, 131.4, 128.8, 128.7, 128.3, 128.1, 127.5, 121.8, 67.2, 53.7, 53.1, 51.7, 41.4, 40.7, 40.2, 25.4–24.8, 23.3–21.5 ppm; HPLC: t_R = 7.16 min, HPLC (intermediate 10a): t_R = 4.98 min; HRMS calcd for [(C₃₅H₅₀N₄O₃)]⁺ m/z : 645.3627, found: 645.3629.

(S)-3-(Benzyloxycarbonyl-(S)-leucyl-(S)-leucylamino)-5-methyl-2-oxo-N-(4-N,N-dimethylaminophenyl)hexanamide (6b): Yield: 16% (20 mg), colorless oil: ¹H NMR (500 MHz, CDCl₃, 300 K): δ = 8.55 (1 H, s), 7.50 (2 H, d, J = 9.0 Hz), 7.33 (5 H, m), 6.87 (1 H, m), 6.75 (2 H, m), 6.46 (1 H, m), 5.39 (1 H, m), 5.23 (1 H, m), 5.10 (2 H, m), 4.49 (1 H, m), 4.18 (1 H, m), 2.94 (6 H, s), 1.65 (6 H, m), 1.50 (3 H, m), 0.90 ppm (18 H, m); ¹³C NMR (125 MHz, CDCl₃, 300 K): δ = 197.0, 172.4, 171.6, 156.4, 156.3, 136.2, 128.8, 128.7, 128.4, 128.2, 121.4, 113.2, 67.4, 53.8, 53.2, 51.8, 41.4, 41.0, 40.7, 40.5, 25.4–24.9, 23.4–21.5 ppm; HPLC: t_R = 7.28 min, HPLC (intermediate 10b): t_R = 5.91 min; HRMS calcd for [(C₃₅H₅₁N₅O₃)]⁺ m/z : 638.3916, found: 638.3917.

(S)-3-(Benzyloxycarbonyl-(S)-leucyl-(S)-leucylamino)-5-methyl-2-oxo-N-(4-methoxyphenyl)hexanamide (6c): Yield: 24% (21 mg), colorless oil: ¹H NMR (500 MHz, CDCl₃, 300 K): δ = 8.26 (1 H, s), 7.55 (2 H, d, J = 9.1 Hz), 7.33 (5 H, m), 6.88 (2 H, d, J = 9.1 Hz), 6.87 (1 H, m), 6.53 (1 H, d, J = 8.2 Hz), 5.38 (1 H, m), 5.29 (1 H, m), 5.10 (2 H, m), 4.49 (1 H, m), 4.19 (1 H, m), 3.79 (3 H, s), 1.66 (6 H, m), 1.49 (3 H, m), 0.92 ppm (18 H, m); ¹³C NMR (125 MHz, CDCl₃, 300 K): δ = 196.9, 172.8, 172.5, 171.8, 157.3, 156.7, 136.2, 130.0, 128.7, 128.4, 128.2, 121.6, 114.5, 67.4, 55.8, 53.8, 53.1, 51.7, 41.3, 40.7, 40.3, 25.4, 24.9, 23.4–21.3 ppm; HPLC: t_R = 7.89 min, HPLC (intermediate 10c): t_R =

6.47 min; HRMS calcd for [(C₃₄H₄₈N₄O₃)]⁺ m/z : 625.3596, found: 625.3611.

(S)-3-(Benzyloxycarbonyl-(S)-leucyl-(S)-leucylamino)-5-methyl-2-oxo-N-(4-cyanophenyl)hexanamide (6d): Yield: 12% (17 mg), colorless oil: ¹H NMR (300 MHz, CDCl₃, 300 K): δ = 8.86 (1 H, s), 7.76 (2 H, m), 7.65 (2 H, m), 7.35 (5 H, m), 6.94 (1 H, m), 6.39 (1 H, m), 5.28 (1 H, m), 5.12 (2 H, m), 4.46 (1 H, m), 4.14 (1 H, m), 1.73 (9 H, m), 0.94 ppm (18 H, m); ¹³C NMR (125 MHz, CDCl₃, 300 K): δ = 196.0, 172.6, 172.1, 157.5, 140.4, 136.0, 133.5, 128.8, 128.6, 128.2, 120.1, 118.6, 108.6, 67.5, 54.0, 52.8, 51.6, 41.1, 40.3, 39.9, 25.4–24.9, 23.3–21.6 ppm; HPLC: t_R = 7.75 min, HPLC (intermediate 10d): t_R = 5.05 min; HRMS calcd for [(C₃₄H₄₆N₄O₃)]⁺ m/z : 620.3448, found: 620.3457.

(S)-3-(Benzyloxycarbonyl-(S)-leucyl-(S)-leucylamino)-5-methyl-2-oxo-N-(4-bromophenyl)hexanamide (6e): Yield: 41% (218 mg), white solid: ¹H NMR (500 MHz, CDCl₃, 300 K): δ = 8.95 (1 H, s), 7.52–7.05 (9 H, m), 5.75 (1 H, s), 5.29 (1 H, m), 5.12–5.01 (2 H, m), 4.60 (1 H, m), 4.31 (1 H, m), 1.57–1.26 (9 H, m), 0.95–0.66 ppm (18 H, m); ¹³C NMR (125 MHz, CDCl₃, 300 K): δ = 196.7, 172.4, 172.1, 157.4, 156.3, 136.2, 135.5, 132.1, 128.5, 128.2, 128.0, 121.5, 118.1, 67.1, 53.6, 53.1, 51.1, 51.4, 41.5, 40.7, 39.8, 25.2, 24.7, 23.7, 23.1, 22.8, 22.5, 22.4, 22.2, 21.4 ppm; HPLC: t_R = 7.68 min, HPLC (intermediate 10e): t_R = 6.05 min; HRMS calcd for [(C₃₃H₄₅BrN₄O₃)]⁺ m/z : 673.2595, found: 673.2605.

Biological and structural analysis

Inhibition assay of purified 20S proteasome: 100 ng of constitutive 20S proteasomes (isolated from human red blood cells) were incubated with defined concentrations of inhibitors 6a–e for 15 min at room temperature. Equal volumes of a protease substrate solution were added (final concentration: 50 μM) and incubated at 37°C for 1 h. Proteasome activity was recorded by the release of the fluorogenic AMC group from the protease substrate at 360 nm excitation and 460 nm emission, (LLE-AMC, VGR-AMC, and LLVY-AMC were used to analyze the different cleavage properties of the proteasome). Data, normalized to controls, represent the means of two independent experiments, each performed in triplicate (n = 2). Statistical analysis was performed following the methods of Cumming et al.^[17]

Intracellular inhibition of proteasomes: HeLa cells (2×10^5 cells per well) were seeded in 96-well microtiter plates and cultured in RPMI supplemented with 10% FCS, 2 mM glutamine, and penicillin-streptomycin (100 U mL⁻¹ penicillin, 100 μg mL⁻¹ streptomycin). Inhibitors were added as a tenfold stock to adjust the indicated concentrations and were incubated overnight at 37°C under 5% CO₂. The supernatant was removed, and cells were washed with cold phosphate-buffered saline and lysed in 100 μL of Tris (20 mM), EDTA (1 mM), and 0.1% NP-40. The proteasome activity was measured in 25 μL of the lysate with a final concentration of 50 μM Suc-LLVY-AMC. The assays were incubated for 1 h at 37°C . Proteasome activity was estimated at 460 nm emission (excitation at 360 nm). Experimental data (see Supporting Information), normalized to controls, represent the means of triplicate experiments (n = 1). Statistical analysis was performed following the methods of Cumming et al.^[17]

Crystallization and structure determination: Crystals of the yeast CP were grown in hanging drops at 20°C as described previously.^[18,19] The protein concentration used for crystallization was 40 mg mL⁻¹ in Tris/HCl (20 mM, pH 7.5) and EDTA (1 mM). Drops contained 1 μL of protein and 1 μL of the reservoir solution (30 mM magnesium

acetate, 100 mM morpholino-ethane-sulfonic acid (pH 7.2), and 10% (w/v) 2-methyl-2,4-pentanediol. Crystals appeared after 2 days and were then soaked with inhibitors in DMSO at final concentrations of 2 mM for at least 24 h. Droplets were then complemented with a cryoprotecting buffer (30% (w/v) 2-methyl-2,4-pentanediol, 20 mM magnesium acetate, 100 mM morpholino-ethane-sulfonic acid, pH 6.9) and supercooled in a stream of liquid nitrogen gas at 100 K (Oxford Cryo Systems). A dataset was collected from the CP-6a complex (PDB code: 4R02) at 2.5 Å and cell parameters of $a=137$ Å, $b=301$ Å, $c=146$ Å, and $\beta=113^\circ$ in the P2₁ space group using synchrotron radiation ($\lambda=1.0$ Å) at the X06SA-beamline (Swiss Light Source, Villigen, Switzerland). X-ray intensities were assessed with the program XDS,^[19] while data reduction was carried out with XSCALE^[19] (Table S11). Electron density was improved by averaging and back-transforming the reflections 10 times over the twofold noncrystallographic symmetry axis using the program package MAIN.^[20] Conventional crystallographic rigid body, positional, and temperature factor refinements were carried out with CNS using the yeast CP structure as a starting model (PDB code: 1RYP) (Table S11), while model building was performed with the program MAIN.

Analysis of the reversibility of proteasome inhibition via dialysis: Red blood cell proteasome solution (500 μ L, 2 μ g proteasome) in 20 mM Tris, pH 7.2, 0.5 mM EDTA, and 0.1 mM acid (TAE) was supplemented with 25 μ L of 0.1 mM inhibitor 5 or 6a, or with 25 μ L of 1% DMSO in water. The final inhibitor concentration was 5 μ M. After pre-incubation for 30 min at room temperature, the samples were transferred to 0.5 mL dialysis tubes and dialyzed against 1 L TAE buffer at room temperature for 72 h. The buffer was exchanged after 6, 22, and 48 h. Duplicates of 25 μ L (50 ng of proteasomes) were used for estimation of activity with 100 μ M of Suc-LLVY-AMC at 460 nm emission (excitation 360 nm) at indicated time points. Experimental data (see Supporting Information) represent the means of duplicate experiments ($n=1$). Statistical analysis was laid out following Cumming et al.^[17]

Molecular docking

Preparation of protein–ligand structures: Only the $\beta 5$ and $\beta 6$ subunits of the complex were prepared; all others were removed from the system. All unbound water molecules were removed from the complexes, because they showed no significant role in ligand binding. Hydrogens were added to the X-ray structures of the CP-5 complex using the Protonate 3D function within MOE2012.10.^[14] The Protonate 3D application assigns protonation states from a discrete collection of states by optimizing the titration free energy of all titratable groups in the context of an all-atom model of a macromolecular structure (including ligands and solvent).^[21] The generalized Born/volume integral electrostatics model is used for longer range interactions and solvation effects.^[22]

Preparation of the ligand database: All five possible *ortho*- and *para*-substituted methyl-derivatives of 5 were derived from the X-ray structure by manipulation with the Builder tool of MOE. For each derivative, the dihedral angle spanning the phenyl moiety and the amide carbonyl was altered in 45° steps using the Builder software, resulting in four different conformations for each C_2 -symmetrical and eight conformations for each unsymmetrical aromatic ring.

Covalent docking of the ligand database: The customized script Conflexdock was used for covalent docking of the ligands.^[23] The acetal carbon connected to Thr10ⁱ was assigned as the “anchor” atom and fixed in position. Conformations from the prepared data-

bases were docked into the active site and energy-minimized using the Amber12EHT force field, which is parameterized for proteins and nucleic acids using Amber and parameterized for small molecules using 2D Extended Hückel Theory.^[24] All non-anchor atoms belonging to the ligand were free to move during the minimization. The pocket atoms of the receptor were tethered to allow movement. All other atoms were fixed. To estimate the binding free energy score, the London dG scoring function was used. For visualization of the procedure, see the Supporting Information.

Supporting Information

Experimental procedures and characterization data for substrate compounds 8a–d and 9a–d; Inhibition of isolated 20S proteasomes by compounds 6a–e; Activity curves for intracellular inhibition of proteasomes by compounds 6a–e; X-ray data collection and refinement statistics; Reversibility of proteasome inhibition by compounds 5 and 6a, as determined by dialysis and proteasome activity; Molecular docking specification; NMR spectra of compounds 6a–e.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), Germany (GRK1657, projects 2E and 3C), and by the Hans und Ilse Breuer Stiftung.

Keywords: 20S proteasome • cancer • drug development • Passerini reaction • ubiquitin

- [1] a) M. Groll, L. Ditzel, J. Lowe, D. Stock, M. Bochtler, H. D. Bartunik, R. Huber, *Nature* **1997**, *386*, 463–471; b) J. Löwe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, R. Huber, *Science* **1995**, *268*, 533–539.
- [2] I. Schechter, A. Berger, *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.
- [3] a) M. Groll, R. Huber, *Int. J. Biochem. Cell Biol.* **2003**, *35*, 606–616; b) A. K. Nussbaum, T. P. Dick, W. Keilholz, M. Schirle, S. Stevanovic, K. Dietz, W. Heinemeyer, M. Groll, D. H. Wolf, R. Huber, H. G. Rammensee, H. Schild, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12504–12509.
- [4] A. F. Kisselev, W. A. van der Linden, H. S. Overkleeft, *Chem. Biol.* **2012**, *19*, 99–115.
- [5] a) D. Chen, M. Frezza, S. Schmitt, J. Kanwar, Q. P. Dou, *Curr. Cancer Drug Targets* **2011**, *11*, 239–253; b) J. L. Thompson, *Ann. Pharmacother.* **2013**, *47*, 56–62.
- [6] A. Asai, T. Tsujita, S. V. Sharma, Y. Yamashita, S. Akinaga, M. Funakoshi, H. Kobayashi, T. Mizukami, *Biochem. Pharmacol.* **2004**, *67*, 227–234.
- [7] M. Groll, O. V. Larionov, R. Huber, A. de Meijere, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4576–4579.
- [8] a) V. S. Korotkov, A. Ludwig, O. V. Larionov, A. V. Lygin, M. Groll, A. de Meijere, *Org. Biomol. Chem.* **2011**, *9*, 7791–7798; b) S. Kawamura, Y. Unno, A. List, A. Mizuno, M. Tanaka, T. Sasaki, M. Arisawa, A. Asai, M. Groll, S. Shuto, *J. Med. Chem.* **2013**, *56*, 3689–3700.
- [9] a) P. G. Richardson, B. Barlogie, J. Berenson, S. Singhal, S. Jagannath, D. Irwin, S. V. Rajkumar, G. Skralovic, M. Alsina, R. Alexanian, D. Siegel, R. Z. Orlowski, D. Kuter, S. A. Limentani, S. Lee, T. Hideshima, D.-L. Esseltine, M. Kauffman, J. Adams, D. P. Schenkein, K. C. Anderson, *New Engl. J. Med.* **2003**, *348*, 2609–2617; b) M. J. Williamson, M. D. Silva, J. Terkelsen, R. Robertson, L. Yu, C. Xia, P. Hatsis, B. Bannerman, T. Babcock, Y. Cao, E. Kupperman, *Mol. Cancer Ther.* **2009**, *8*, 3234–3243.
- [10] a) H. A. Braun, S. Umbreen, M. Groll, U. Kuckelkorn, I. Mlynarczyk, M. E. Wigand, I. Drung, P.-M. Kloetzel, B. Schmidt, *J. Biol. Chem.* **2005**, *280*, 28394–28401; b) S. Chatterjee, D. Dunn, S. Maliya, M. A. Ator, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2603–2606; c) M. L. Stein, H. Cui, P. Beck, C. Dubiella, C. Voss, A. Krüger, B. Schmidt, M. Groll, *Angew. Chem. Int. Ed.* **2014**, *53*, 1679–1683; *Angew. Chem.* **2014**, *126*, 1705–1709.

- [11] K. P. Dhake, P. J. Tambade, R. S. Singhal, B. M. Bhanage, *Green Chem. Lett. Rev.* **2011**, *4*, 151–157.
- [12] R. B. Lacerda, C. K. de Lima, L. L. da Silva, N. C. Romeiro, A. L. Miranda, E. J. Barreiro, C. A. Fraga, *Bioorg. Med. Chem.* **2009**, *17*, 74–84.
- [13] a) L. P. Hammett, *J. Am. Chem. Soc.* **1937**, *59*, 96–103; b) J. Shorter, *Chem. Unserer Zeit* **1985**, *19*, 197–208; c) T. Yoshida, M. Shimizu, M. Harada, S. Hitaoka, H. Chuman, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 124–128.
- [14] Molecular Operating Environment (MOE) 2012.10, Chemical Computing Group Inc., 1010 Sherbooke Street West, Suite #910, Montreal, QC, H3A 2R7 (Canada), **2012**.
- [15] H. Kubinyi in *Pharmacokinetic Optimization in Drug Research*, Helvetica Chimica Acta, Zürich, **2007**, pp. 513–524.
- [16] M. Frigerio, M. Santagostino, S. Sputore, *J. Org. Chem.* **1999**, *64*, 4537–4538.
- [17] G. Cumming, F. Fidler, D. L. Vaux, *J. Cell Biol.* **2007**, *177*, 7–11.
- [18] M. Groll, R. Huber, *Methods Enzymol.* **2005**, *398*, 329–336.
- [19] W. Kabsch, *J. Appl. Crystallogr.* **1993**, *26*, 795–800.
- [20] D. Turk, PhD thesis, Technische Universität München, **1992**.
- [21] P. Labute, *Proteins Struct. Funct. Bioinf.* **2009**, *75*, 187–205.
- [22] P. Labute, *J. Comput. Chem.* **2008**, *29*, 1693–1698.
- [23] Scientific Vector Language (SVL) source code provided by Chemical Computing Group Inc., 1010 Sherbooke Street West, Suite #910, Montreal, QC, H3A 2R7 (Canada), **2012**.
- [24] R. Salomon-Ferrer, D. A. Case, R. C. Walker, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 198–210.

Received: June 12, 2014
Published online on August 1, 2014

Supporting Information

© Copyright Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, 2014

α -Keto Phenylamides as P1'-Extended Proteasome Inhibitors

Constantin Voss,^[a] Christoph Scholz,^[a] Sabine Knorr,^[b] Philipp Beck,^[c] Martin L. Stein,^[c]
Andrea Zall,^[a] Ulrike Kuckelkorn,^[d] Peter-Michael Kloetzel,^[d] Michael Groll,^[c] Kay Hamacher,^[b]
and Boris Schmidt^{*,[a]}

cmdc_201402244_sm_miscellaneous_information.pdf

Table of contents:

S2	Experimental procedures and characterization data for substrate compounds 8a-8d and 9a-9d
S4	Inhibition of isolated 20S proteasomes by compounds 6a-6e
S5	Activity curves for intracellular inhibition of proteasomes by compounds 6a-6e
S6	X-ray data collection and refinement statistics
S9	Reversibility of proteasome inhibition by compounds 5 and 6a determined by dialysis and proteasome activity
S10	Molecular docking specification
S11	NMR spectra of compounds 6a-6e
S22	Reference list

Experimental procedures and characterization data for the substrate compounds 8a-8d and 9a-9d

Synthesis of phenyl formamides from anilines (8a-8d). Aniline (1.0 eq.) and HCOOH (2.4 eq.) are stirred in a capped vial and heated to 60°C for 15 h. After the addition of CH₂Cl₂, the mixture is washed with water (3x), saturated aqueous NaHCO₃ (3x) and brine (3x). The organic layer is dried over Na₂SO₄, filtered and concentrated under reduced pressure.

N-(2,4-dimethylphenyl)formamide (8a): Yield: 97%, brown solid. ¹H-NMR (300 MHz, MeOD, 300 K): δ [ppm] = 8.45 (1H_a, d, *J* = 11.3 Hz), 8.38 (1H_b, d, *J* = 1.8 Hz), 8.23 (1H_a, m), 7.66 (1H_b, d, *J* = 8.7 Hz), 7.36 (1H_b, m), 7.01 (3H, m), 2.31 (3H_a, s), 2.28 (3H_b, s), 2.27 (3H_a, s), 2.22 (3H_b, s). Trans-rotamer: a (60%), cis-rotamer b (40%). HPLC: t_R = 3.14 min.

N-(4-(dimethylamino)phenyl)formamide (8b): Yield: 56%, black solid. ¹H-NMR (300 MHz, CDCl₃, 300 K): δ [ppm] = 8.45 (1H_a, d, *J* = 11.6 Hz), 8.26 (1H_b, d, *J* = 1.8 Hz), 8.09 (1H_b, m), 7.55 (1H_a, m), 7.38 (2H_b, *J* = 9.0 Hz), 6.98 (2H_a, d, *J* = 9.0 Hz), 6.69 (2H_{a+b}, d, *J* = 9.0 Hz), 2.93 (3H_a, s), 2.91 (3H_b, s). Trans-rotamer: a (50%), cis-rotamer b (50%). HPLC: t_R = 0.76 min.

N-(4-methoxyphenyl)formamide (8c): Yield: 74%, violette solid. ¹H-NMR (300 MHz, CDCl₃, 300 K): δ [ppm] = 8.51 (1H_a, d, *J* = 11.4 Hz), 8.44 (1H_b, m), 8.29 (1H_b, d, *J* = 1.9 Hz), 7.74 (1H_a, m), 7.44 (2H_b, d, *J* = 9.1 Hz), 7.03 (2H_a, d, *J* = 9.0 Hz), 6.87 (2H_a, d, *J* = 9.0 Hz), 6.84 (2H_b, d, *J* = 9.0 Hz), 3.79 (3H_a, s), 3.77 (3H_b, s). Trans-rotamer: a (50%), cis-rotamer b (50%). HPLC: t_R = 1.96 min.

N-(4-cyanophenyl)formamide (8d): Yield: 99%, white solid. ¹H-NMR (300 MHz, MeOD, 300 K): δ [ppm] = 8.79 (1H_a, s), 8.40 (1H_b, s), 7.82 (2H_b, d, *J* = 8.5 Hz), 7.72 (2H_b, 2H_a, m), 7.38 (2H_a, d, *J* = 8.5 Hz). Trans-rotamer: a (20%), cis-rotamer b (80%). HPLC: t_R = 2.09 min.

Synthesis of phenyl isonitriles from phenyl formamides (9a-9d). Phenyl formamide (1.0 eq.) and NEt₃ (10.0 eq.) are dissolved in dry CH₂Cl₂ and cooled to 0°C. After the addition of POCl₃ (2.1 eq.) are added and the mixture is allowed to warm slowly to room temperature and stirred for 12 h. Due to the sensibility of the product to hydrolysis, the work up has to be done fast with the use of cooled solvents! After addition of CH₂Cl₂ the mixture is washed with saturated aqueous NaHCO₃ (2x). The organic layer is then dried over Na₂SO₄, filtered and concentrated under reduced pressure at room temperature. Purification can be done by a fast liquid chromatography (for solids) or carefully via Kugelrohr-distillation (for liquids). Isonitriles should be converted instantly due to the fast hydrolysis on air. All operations should be done in a well-ventilated hood due to the strong smell and high toxicity of isonitrile compounds.

1-isocyano-2,4-dimethylbenzene (9a): Yield: 61%, colorless liquid. HPLC: t_R = 6.48 min. **9a** was directly converted in the next reaction without further characterization.

4-isocyano-N,N-dimethylaniline (9b): Yield: 90%, brown solid. HPLC: t_R = 4.87 min. **9b** was directly converted in the next reaction without further characterization.

1-isocyano-4-methoxybenzene (9c): Yield: 90%, brown solid. HPLC: t_R = 5.31 min. **9c** was directly converted in the next reaction without further characterization.

4-isocyanobenzonitrile (9d): Yield: 53%, white solid. HPLC: t_R = 4.49 min. **9d** was directly converted in the next reaction without further characterization.

Inhibition of isolated 20S proteasomes by compounds 6a-6e

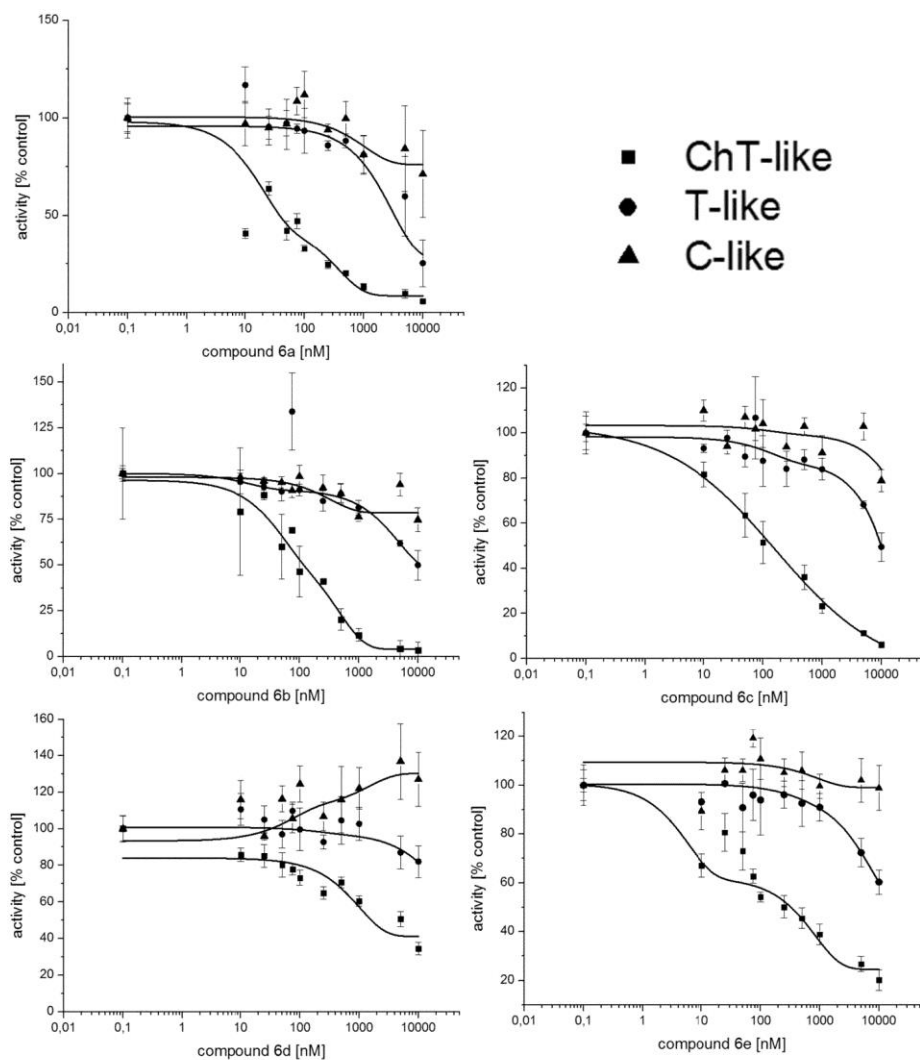


Figure S 1: Activity curves of constitutive 20S proteasomes (isolated from human red blood cells) by compounds **6a-6e**. 100 ng of 20S proteasomes were incubated with defined

S4

concentrations of inhibitors **6a-6e** for 15 min at room temperature, respectively. Equal volumes of a protease substrate solution were added (final concentration 50 μ M) and incubated at 37°C for 1 h. The proteasome activity was evaluated by the release of the fluorogenic AMC group from the protease substrate at 360 nm excitation and at 460 nm emission, (LLE-AMC, VGR-AMC and LLVY-AMC were used to analyze the different cleavage properties of the proteasome). Data normalized to controls represent the mean of two independent experiments, each performed in triplicates ($n = 2$). Error bars show the standard deviation (SD). Statistical analysis was laid out following Cumming et al..¹

Activity curves for intracellular inhibition of proteasomes by compounds **6a-6e**

In vivo (ChT-like):

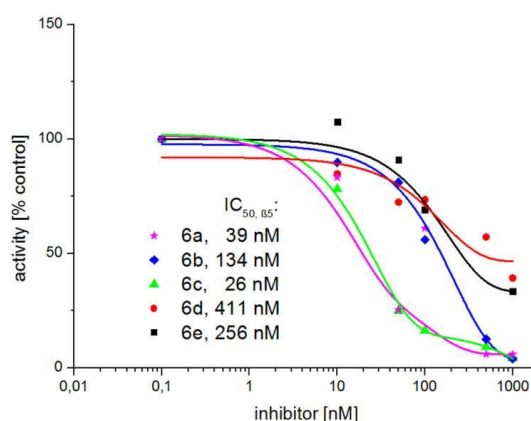


Figure S 2: Activity curves for intracellular inhibition of proteasomes by compounds **6a-6e**. 2×10^5 HeLa cells were seeded in 96 well microtiter plates and cultured in RPMI supplemented with 10% FCS, 2 mM glutamine and Penicillin/Streptomycin. Inhibitors were added as 10-fold stock to adjust the indicated concentrations and were incubated over night at 37°C under 5%

S5

CO₂. The supernatant was removed, cells were washed with cold phosphate buffered saline and lysed in 100 µl 20 mM Tris, 1 mM EDTA and 0,1% NP-40. The proteasome activity was measured in 25 µl of the lysats with final concentrations of 50µM Suc-LLVY-AMC. The assays were incubated for 1 hour at 37°C. Proteasome activity was estimated at 460 nm emission (excitation at 360 nm). Normalized data represent means of triplicate experiments to indicate cell permeability (n = 1). Statistical analysis was laid out following Cumming et al..¹

X-ray data collection and refinement statistics

Table ST1. Crystallographic data collection and refinement statistics of the CP in complex with 6a (BSc4999).

Crystallographic data	CP: 6a *
Crystal parameters	
Space group	P2 ₁
Cell constants (Å/°)	a = 137.2
(dataset was collected from	b = 301.3
1 crystal / 1 CP per AU	c = 146.3
	β = 113.0
Data collection	
Beamline	X06SA, SLS
Wavelength, Å	1.0
Resolution range, Å [†]	30-2.5
No. observations	1093674
No. unique reflections [‡]	363107
Completeness, % [†]	96.9 (99.1)
R _{merge} ^{†,§}	7.4 (55.8)
I/σ (I) [†]	10.9 (2.0)

Refinement		
Resolution range, Å	15-2.5	
No. reflections working set	344951	
No. reflections test set	18156	
No. nonhydrogen	49376	
No. of ligand atoms	90	
Water/Magnesium Atoms	2707/11	
R _{work} /R _{free} % [†]	17.6/20.7	
Rmsd bond (Å)/(°)**	0.004/0.968	
Ramachandran plot, %***	97.4/2.2/0.4	

*Dataset has been collected on a single crystal.

†Values in parentheses of resolution range, completeness, Rmerge, and I/σ (I) correspond to the last resolution shell.

‡Friedel pairs were treated as identical reflections.

§ $R_{\text{merge}}(I) = \frac{\sum_{\text{hkl}} \sum_j |I(\text{hkl})_j - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} I(\text{hkl})}$, where $I(\text{hkl})_j$ is the measurement of the intensity of reflection hkl and $\langle I(\text{hkl}) \rangle$ is the average intensity.


¶ $R = \frac{\sum_{\text{hkl}} |F_{\text{obs}}| - |F_{\text{calc}}|}{\sum_{\text{hkl}} |F_{\text{obs}}|}$, where R_{free} is calculated without a sigma cut off for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections.

**Deviations from ideal bond lengths/angles.

***Number of residues in favored region/allowed region/outlier region.

Crystallization and soaking experiments. Crystals of the yeast CP were grown in hanging drops at 20°C as described previously.^{2, 3} The protein concentration used for crystallization was 40 mg/mL in Tris / HCl (20 mM, pH 7.5) and EDTA (1 mM). The drops contained 1 µL of protein and 1 µL of the reservoir solution (30 mM magnesium acetate, 100 mM morpholino-ethane-sulfonic acid (pH 7.2) and 10% (wt/vol) 2-methyl-2,4-pentanediol). Crystals appeared after two days and were then soaked with inhibitors in DMSO at final concentrations of 2 mM for at least 24 hours. Droplets were then complemented with a cryoprotecting buffer (30% (wt/vol) 2-methyl-2,4-pentanediol, 20 mM magnesium acetate, 100 mM morpholino-ethane-sulfonic acid, pH 6.9) and supercooled in a stream of liquid nitrogen gas at 100 K (Oxford Cryo Systems). A dataset was collected from the CP:6a complex (PDB code: 4R02) at 2.5 Å and cell parameters of $a = 137$ Å, $b = 301$ Å, $c = 146$ Å, and $\beta = 113^\circ$ in the space group $P2_1$ by using synchrotron radiation ($\lambda = 1.0$ Å) at the X06SA-beamline (Swiss Light Source, Villigen, Switzerland). X-ray intensities were assessed with the programme XDS,⁴ while data reduction

S7



was carried out with XSCALE⁴ (Table ST1). Electron density was improved by averaging and back-transforming the reflections 10 times over the twofold noncrystallographic symmetry axis using the program package MAIN.⁵ Conventional crystallographic rigid body, positional, and temperature factor refinements were carried out with CNS using the yeast CP structure as starting model (PDB accession code 1RYP) (Table ST1), while model building was performed with the program MAIN.

Reversibility of proteasome inhibition by compounds **5 and **6a** determined by dialysis and proteasome activity**

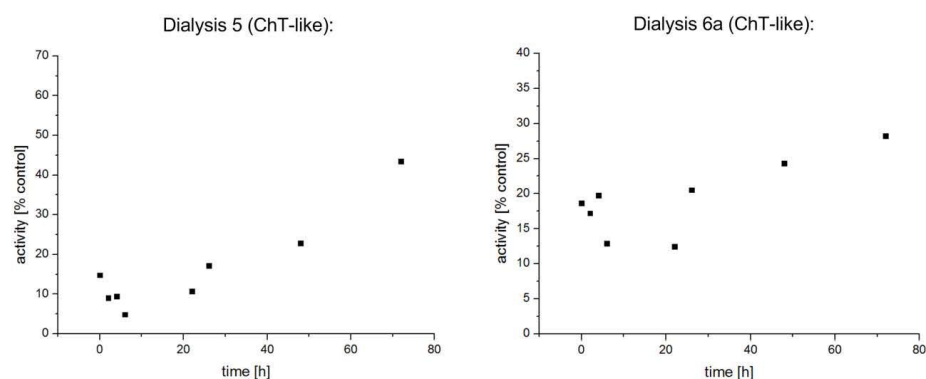


Figure S 3: Analysis of the reversibility of proteasome inhibition by compounds **5** and **6a** via dialysis. 500 μ L of red blood cell proteasome solution (2 μ g proteasome) in 20 mM Tris, pH 7.2, 0.5 mM EDTA, 0.1 mM acid (TAE) was supplemented with 25 μ l of 0.1 mM inhibitor **5** and **6a**, respectively, or with 25 μ l of 1% DMSO in water. Final inhibitor concentration was 5 μ M. After pre-incubation for 30 min at room temperature the samples were transferred to 0.5 ml dialysis tubes and dialyzed against 1 L TAE buffer at room temperature for 72 h. The buffer was exchanged after 6, 22 and 48 h. Duplicates of 25 μ l (50 ng of proteasomes) were used for estimation of activity with 100 μ M of Suc-LLVY-AMC at 460 nm emission (excitation 360 nm) at indicated time points. Each data point represents the mean of a duplicate experiment to indicate a reversible binding mode to the proteasome ($n = 1$). Statistical analysis was laid out following Cumming et al.¹

Molecular docking specification

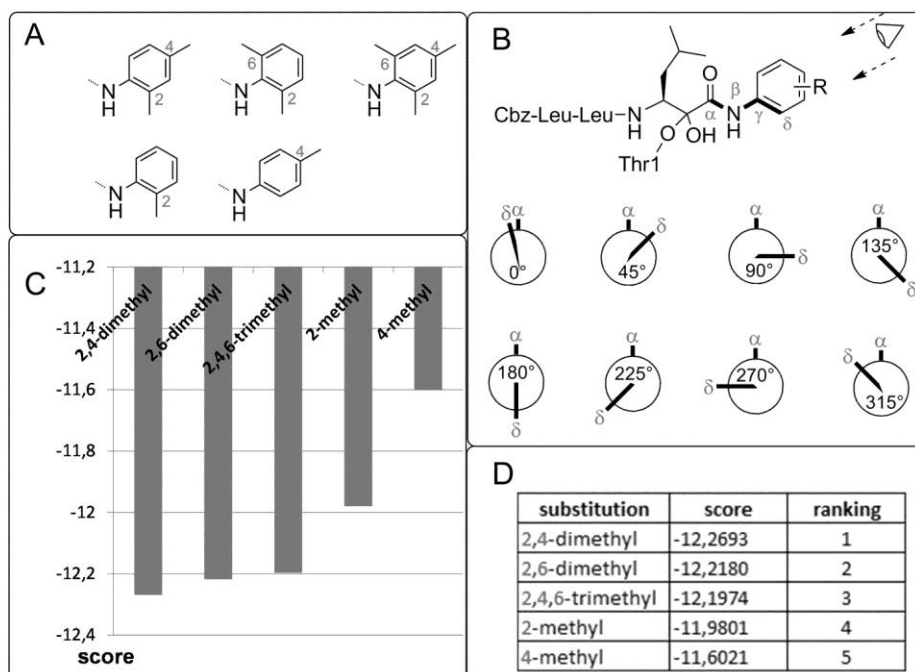


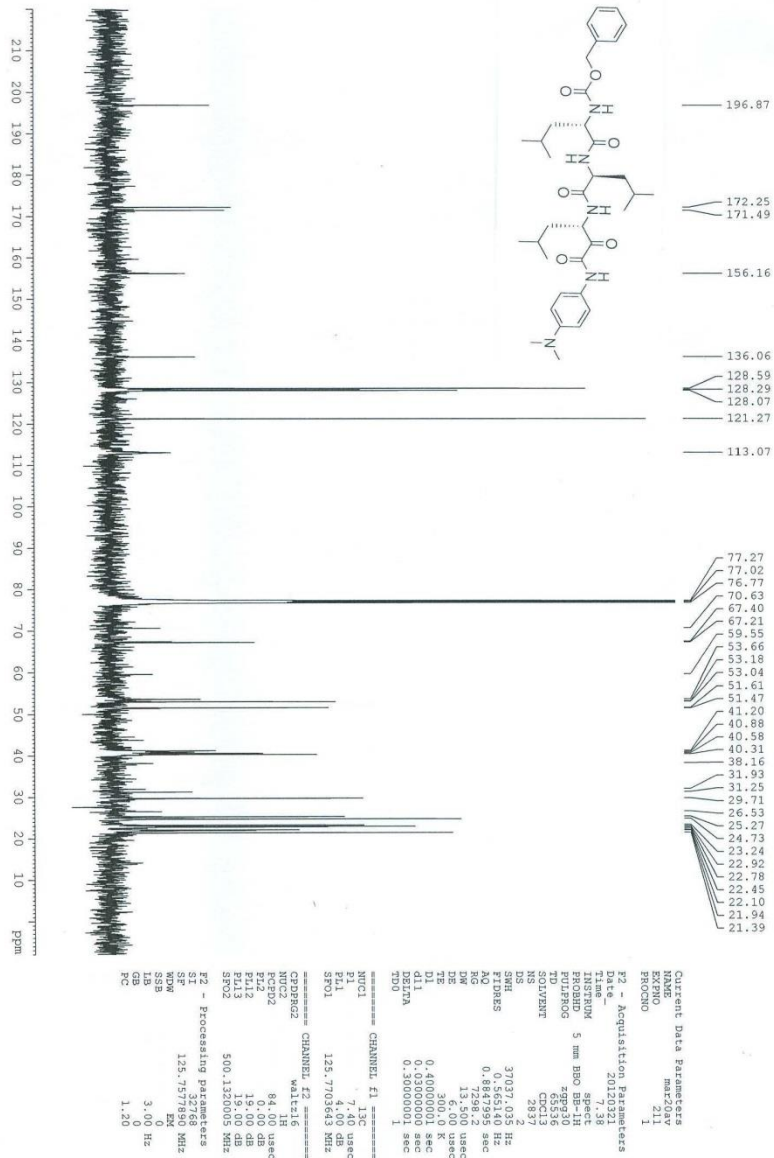
Figure S 4: Visualization of the molecular docking procedure for phenylamide moieties. A: The five possible *ortho*- and *para*-methylphenylamide substitutions. B: The eight starting points for energy minimizations from different spanned dihedral angles for each substitution. The lowest relative value after energy minimizations represents the lowest energy minimum for the respective substitution and is called score. C: Ranked scores for the five substitutions in a bar chart. D: Table of values. Values are dimensionless and can be considered as relative rankings based on energy minima.



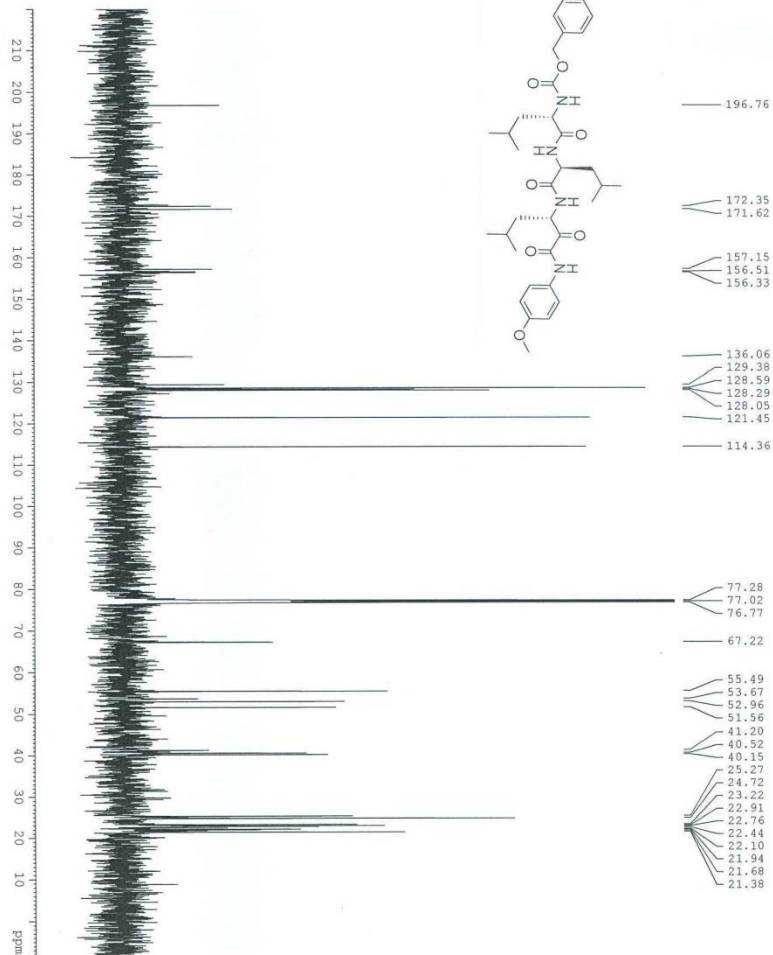
NMR spectra of compounds 6a-6e

S11









Current Data Parameters

NAME: mar21av

EXPNO: 60

PROCNO: 1

F2 - Acquisition Parameters

Time: 02:23.31

Time-: 10.44

INSTRUM: 5 mm BBO spect

PULPROG: zgpg30

TD: 65536

DELTA: 0.000001 sec

SOLVENT: CDCl3

DS: 2

SWH: 37037.015 Hz

FIDRES: 0.133 Hz

AQRES: 0.887995 sec

RG: 14596.5

RG: 13.500 usec

DM: 13.500 usec

TE: 300.2 K

DE: 0.000001 sec

DELTA: 0.000001 sec

TD0: 1

===== CHANNEL f1 =====

NUC1: 13C

P1: 7.40 usec

PL1: 0.00 dB

SFO1: 125.7703643 MHz

===== CHANNEL f2 =====

NAME: mar216

EXPNO: 1H

PROCNO: 1H

F2 - Acquisition Parameters

Time: 02:23.31

Time-: 10.44

INSTRUM: 5 mm BBO spect

PULPROG: zgpg30

TD: 65536

DELTA: 0.000001 sec

SOLVENT: CDCl3

DS: 2

SWH: 37037.015 Hz

FIDRES: 0.133 Hz

AQRES: 0.887995 sec

RG: 14596.5

RG: 13.500 usec

DM: 13.500 usec

TE: 300.2 K

DE: 0.000001 sec

DELTA: 0.000001 sec

TD0: 1

===== CHANNEL f3 =====

NAME: mar216

EXPNO: 1H

PROCNO: 1H

F2 - Acquisition Parameters

Time: 02:23.31

Time-: 10.44

INSTRUM: 5 mm BBO spect

PULPROG: zgpg30

TD: 65536

DELTA: 0.000001 sec

SOLVENT: CDCl3

DS: 2

SWH: 37037.015 Hz

FIDRES: 0.133 Hz

AQRES: 0.887995 sec

RG: 14596.5

RG: 13.500 usec

DM: 13.500 usec

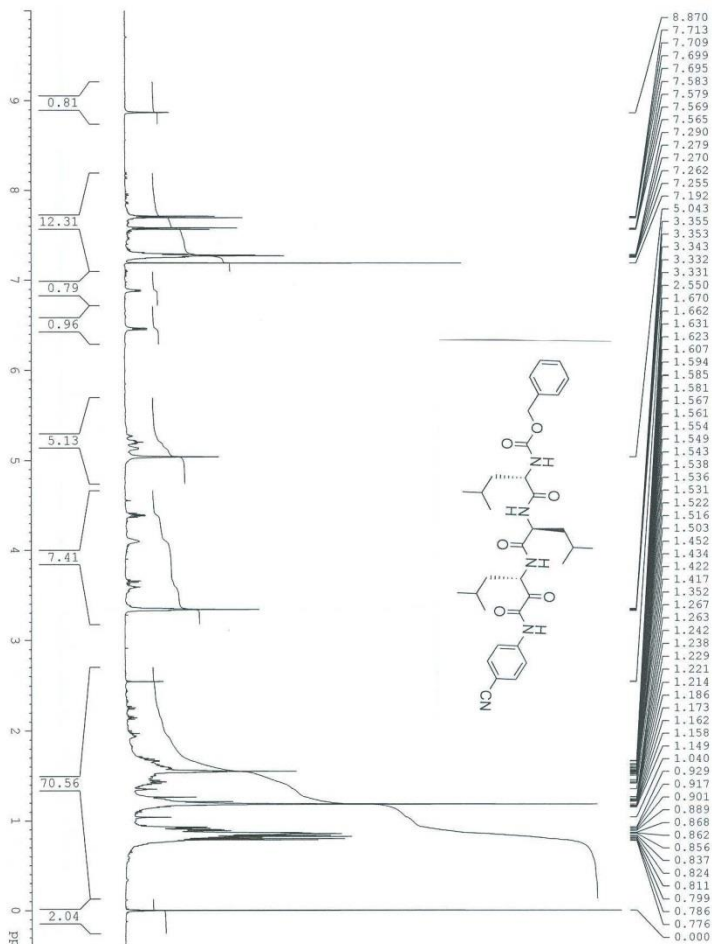
TE: 300.2 K

DE: 0.000001 sec

DELTA: 0.000001 sec

TD0: 1

Scholz 055

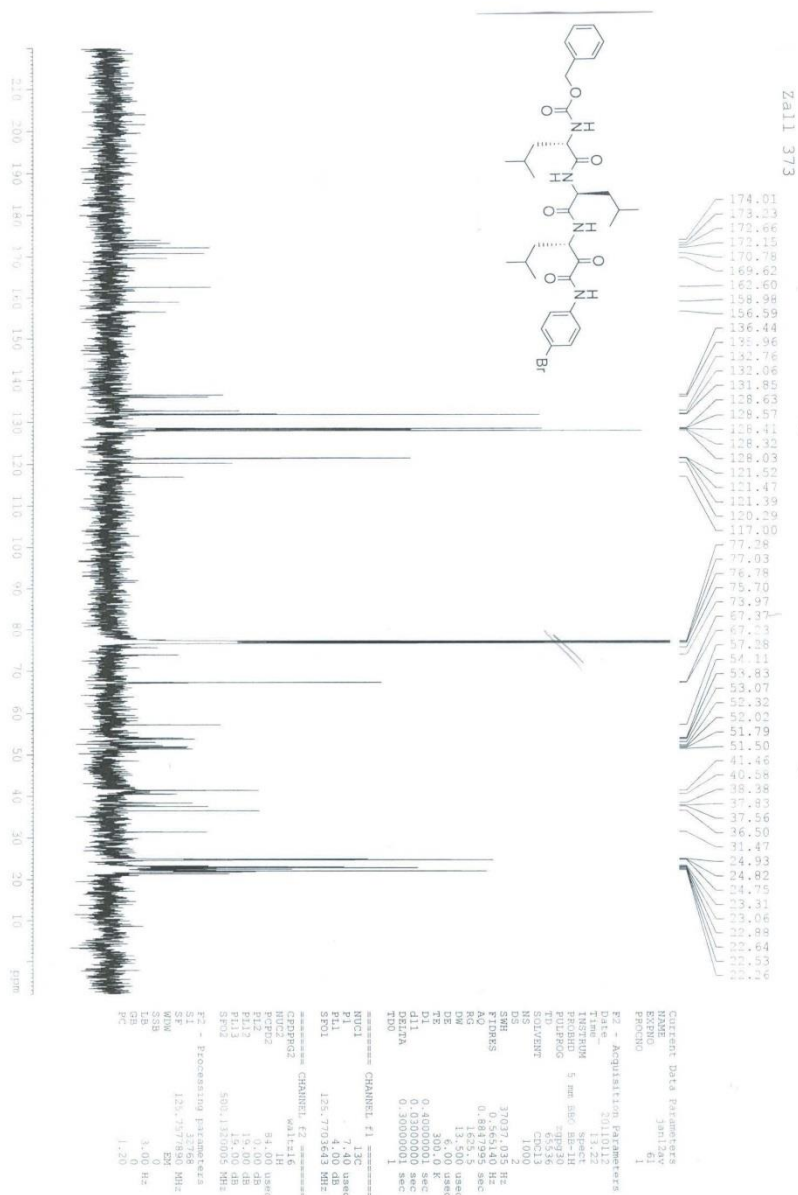


Current Data Parameters
 EXNO 60
 PROCNO 1
 F2 - Acquisition Parameters
 Date_ 20120523
 Time_ 13.36
 INSTRUM spect
 PROBD 5 mm BBO BB-1H
 PULPROG zgpg30
 PC 64
 SOLVENT CDCl3
 NS 64
 DS 4
 SWH 10390.578 Hz
 FIDRES 0.157632 Hz
 AQ 3.1172632 sec
 RG 256
 DW 48.400 usec
 DE 3.00 usec
 TE 300.0 K
 D1 0.50000000 sec
 TDO 1
 CHANNEL F1
 NUC1 1H
 P1 11.10 usec
 PL 0.00 dB
 SFO1 500.130885 MHz
 F2 - Processing parameters
 SI 32768
 SF 500.130885 MHz
 WF 0
 SSF 0.10 Hz
 GB 0
 PC 1.00

S18



72



S21

-
- (1) Cumming, G.; Fidler, F.; Vaux, D. L. Error bars in experimental biology. *The Journal of Cell Biology* **2007**, *177*, 7-11.
 - (2) Groll, M.; Ditzel, L.; Löwe, J.; Stock, D.; Bochtler, M.; Bartunik, H.; Huber, R. Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **1997**, *386*, 463-71.
 - (3) Groll, M.; Huber, R. Purification, crystallization, and X-ray analysis of the yeast 20S proteasome. *Methods Enzymol* **2005**, *398*, 329-36.
 - (4) Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 125-32.
 - (5) Turk, D. Improvement of a Program for Molecular Graphics and Manipulation of Electron Densities and Its Application for Protein Structure Determination **1992**, *PhD thesis*, Technische Universität München.

3.2 Untersuchung des Einflusses eines Ligand/Met45-Kontaktes auf die enzymatische Aktivität des Proteasoms

Der Inhalt dieses Kapitels wird voraussichtlich 2015 zur Veröffentlichung eingereicht:

Christoph Scholz, Constantin Voss, Sabine Knorr, Ulrike Kuckelkorn, Kay Hamacher, Peter M. Klotzel, Boris Schmidt, "Paradigm Caught by a Mouse Trap: 20S Proteasome's $\beta 5$ Subunit is Not Chymotrypsin-like."

Röntgenkristallographische Untersuchungen des Proteasoms im Komplex mit Inhibitoren geben als Momentaufnahmen nur bedingt Einblicke in die dynamischen Prozesse der Ligandenbindung. Man kann zwar an ihnen ablesen, dass es durch Bindung von Bortezomib **5** im Vergleich zum nativen Proteasom zu einer Konformationsänderung von Met45 kommt und sich die S1-Tasche somit perfekt an den P1-Rest des Liganden anpasst, jedoch nicht, ob es während dieses *induced-fit* zu weiteren dynamischen Effekten, wie einer Änderung des H-Brücken-Netzwerkes kurz vor der Bindung, kommt. Zu Beginn dieser Arbeit sollte ein solcher dynamischer „Mausefallenartiger“ Mechanismus untersucht werden, demzufolge es durch Kontakt des P1-Restes eines Liganden („Maus“) mit der Met45-Seitenkette („Köder“) zum Auslösen einer kinetischen Aktivierung des nukleophilen Thr10 γ kommen würde.

Um diesen postulierten Mechanismus zu untersuchen, wurden verschiedene Ansätze verfolgt. Zunächst wurden Derivate des aldehydischen Inhibitors BSc2118 **3** mit unterschiedlich langen P1-Resten (Methyl, Ethyl und Propyl) synthetisiert und *in vitro* an isoliertem Proteasom auf ihre inhibitorische Aktivität hin untersucht. Hierbei war ein deutlicher Sprung in der Aktivität zwischen P1 = Methyl und P1 = Ethyl zu erkennen, welcher durch zeitgleich durchgeführte thermodynamische Berechnungen der relativen freien Bindungsenthalpien nicht zu erklären war. Der beobachtete Effekt musste demnach kinetischer Natur sein und konnte anhand des postulierten *induced-fit*-Mechanismus, der augenscheinlich ab einer Länge von P1 = Ethyl ausgelöst wird, erklärt werden. Eine Moleküldynamiksimulation über einen Zeitraum von 50 ns (siehe Abschnitt 1.2) unterstützte diesen Befund, indem zwei stabile Zustände der Met45-Seitenkette identifiziert werden konnten, die als aktivierter und deaktivierter Zustand der Mausefalle interpretiert wurden.

Da es sich bei dem postulierten Mechanismus um eine kinetische Aktivierung handelt, war der logische Folgeschritt eine kinetische Analyse zur Bestimmung der Parameter k_{cat} , K_M und V_{max} in Abhängigkeit der Länge der P1-Reste geeigneter Substrate. Hierfür wurden fluorogene Substrate, strukturanalog zu den untersuchten Inhibitoren, mit einer AMC 12-Kopfgruppe hergestellt. Werden diese Substrate durch das Proteasom hydrolysiert, setzen sie das fluoreszente AMC 12 frei (siehe Abschnitt 1.3.3), dessen spektroskopische Detektion direkt Rückschlüsse auf die Geschwindigkeit des Substratumsatzes erlaubt. Die Auswertung erfolgte klassisch über die MICHAELIS-MENTEN-Kinetik und einen doppelt reziproken LINEWEAVER-BURK-Plot.

Die MICHAELIS-MENTEN-Konstante K_M , welche als Affinität eines Substrates zum aktiven Zentrum interpretiert werden kann, lag für alle Substrate in einem sehr ähnlichen Bereich. Dies lässt sich durch die Flexibilität der Met45-Seitenkette und den damit einhergehenden *induced-fit* erklären. Der Trend der kinetischen Parameter k_{cat} und V_{max} jedoch stand im starken Kontrast zu den vorherigen Ergebnissen und widerlegte eine kinetische Aktivierung durch große P1-Reste und damit den postulierten „Mausefallen-Mechanismus“. Substrate mit kleinen P1-Resten (Methyl und Ethyl) wurden fast doppelt so schnell abgebaut wie das Propylsubstrat und das etablierte Referenzsubstrat suc-LLVY-AMC mit Tyrosin als P1-Rest. Somit entspricht die Substratpräferenz der $\beta 5$ -Untereinheit eher einer Elastase-ähnlichen oder SNAAP-Aktivität, statt einer ChTL-Aktivität und sollte reklassifiziert werden.

In einem anschließenden *Knockout*-Experiment, bei dem die $\beta 5$ -Untereinheit durch Inkubation mit einer hohen Konzentration Inhibitor ausgeschaltet wurde, konnte eine weitere wichtige Erkenntnis gewonnen werden. Demnach wird das weitläufig als fluorogenes Referenzsubstrat für die $\beta 5$ -Untereinheit eingesetzte suc-LLVY-AMC noch von einem weiteren aktiven Zentrum des Proteasoms hydrolysiert. Es wurde eine Restaktivität des Substratumsatzes von 12% ermittelt, wohingegen sie für die Substrate mit kleineren P1-Resten nur bei 0,5% lag. Die weit verbreitete Anwendung von suc-LLVY-AMC, welches als gängiges Referenzsubstrat abgelöst werden sollte, ist somit besonders im Hinblick auf Selektivitätsmessungen problematisch.

Beitrag C. Scholz: Verfassen des Manuskriptes, Synthese der Inhibitoren 1-4, Synthese des Substrates 7, Berechnung der relativen freien Bindungsenthalpien, Auswertung der Assay-Ergebnisse (Nummerierung laut Manuskript).

Paradigm Caught by a Mouse Trap. 20S Proteasome's $\beta 5$ Subunit is Not Chymotrypsin-like.

*Christoph Scholz, Constantin Voss, Sabine Knorr, Ulrike Kuckelkorn, Kay Hamacher, Peter-Michael Klotzel and Boris Schmidt**

The proteome, the entirety of proteins in an organism at a given time, underlies a complex and dynamic equilibrium between protein synthesis and degradation. In eukaryotes, the ubiquitin/proteasome system (UPS) is the major ATP-dependent degradation pathway. Ubiquitin, a small 8 kDa protein, is covalently attached to a substrate, which is subsequently recognized and degraded to small fragments by the 26S proteasome.^[1] This 26S proteasome is a 2.5 MDa multicatalytical threonine protease complex that consists of two regulatory 19S subunits, which recognize and unfold ubiquitylated substrates, and a cylinder-shaped core particle (CP), the 20S proteasome, which harbours the active sites. The CP consists of 28 subunits with an $\alpha_7\beta_7\beta_7\alpha_7$ symmetry and features six active sites bearing three different substrate specificities: Classified as caspase-like activity (C-like, $\beta 1$) because of its tendency to cleave peptides on the carboxyl side of glutamyl or aspartyl sidechains, the trypsin-like activity (T-like, $\beta 2$) due to the cleavage after basic P1 residues and the chymotrypsin-like activity (ChT-like, $\beta 5$) which predominantly attacks the carboxyl side of neutral, hydrophobic residues.^[2] The observed specificities arise from the unique shape of the substrate binding pockets which are named S1-S4 in C-terminal and S1'-S4' in N-terminal direction of a bound substrate, emanating from its cleavage position which is the alcohol oxygen of Thr¹ (Thr¹O^y). Ligand parts reaching the substrate binding pockets follow a similar nomenclature where S is substituted by P, respectively.

The UPS is involved in several physiologically important cellular processes such as cell cycle progression, immune response or signal transduction, which makes the CP an attractive therapeutic target for cancer and autoimmune disorders.^[3] In fact, the dipeptidic boronic acid bortezomib **S1** and the tetrapeptidic epoxyketone carfilzomib **S2** are currently prescriptive drugs against various types of cancer, in particular relapsed multiple myeloma and mantle cell lymphoma.^[4] Particularly the $\beta 5$ subunit is in the focus of therapeutic targeting, because its inhibition leads to high rates of apoptosis in tumor cells, which is accompanied by a protective effect on healthy cells.^[5] In 2005 we identified the tripeptidic aldehyde BSc2118 (**1**; Scheme 1 a) as a very potent ($IC_{50,\beta 5} = 58$ nM) and specific covalent CP

inhibitor.^[6] The special feature of **1** is a sterically demanding *tert*.-butyl aspartate P2 residue, which discriminates inhibition against the majority of the protease environment. In a recent study **1** showed a cytotoxic effect against human tumor cells. In fact, the determined therapeutic index (LD₅/ED₉₅) of **1** (26.5) in the cellular model is significantly higher than that of bortezomib **S1** (18.4), hence **1** was highly recommended for further preclinical evaluation.^[7]

Besides the proteasome's role in tumorigenesis, peptide fragments produced by the UPS can be presented by major histocompatibility complex (MHC) class I molecules, which makes the proteasome to a fundamental part of the adaptive immune system.^[8] Thus a detailed knowledge of the different substrate cleavage preferences and kinetics of respective proteasome subunits are of high interest not only for immunologists. This resulted in the development of several *in silico* methods for fragment prediction and shed light on the kinetics and mechanisms underlying the UPS-mediated degradation of proteins. These tools furthered the understanding of how different factors, such as the length of a substrate or the presence of regulatory molecules, influence proteolysis.^[9] However, due to the manifold events that occur inside the proteasome, some of them are not conclusively understood today, the prediction of these enigmatic cleavage patterns is still prone to errors.

In this work, we initially investigated a novel kinetically controlled substrate recognition mechanism, which could provide another fundamental piece of the UPS puzzle. Our hypothesis was that a kinetic acceleration of the hydrolytic cleavage arises from hydrophobic interactions between a substrate's P1 residue and the highly flexible sidechain of Met⁴⁵ within the proteasome's $\beta 5$ subunit. This extraordinary flexibility is well established.^[2] For example, in the bortezomib/proteasome complex (PDB-ID: 2F16) Met⁴⁵ shows a dihedral angle α of -174° compared to +48° in the native protein (PDB-ID: 1RYP).^[10] A concerted mechanism may accelerate the binding of substrates with large P1 residues due to a widening of the dihedral angle α of Met⁴⁵. This induced fit was thought to imply a conformational change of the S1 subpocket's backbone which would strengthen the coordination of a catalytically active water molecule through a network of hydrogen bonds and increase the nucleophilicity of Thr¹O^y. An illustrative analogy for this triggered activation is the release of a mouse trap after the mouse touches the lure (Scheme 1 b).

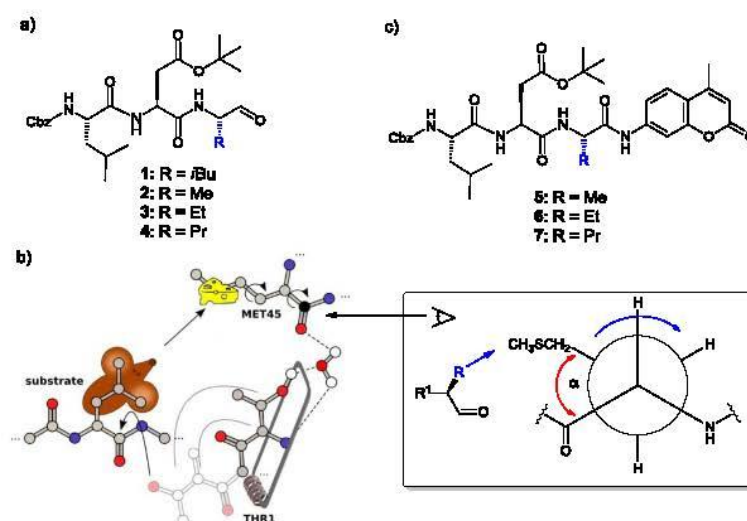


Figure 1: a) Chemical structures of the tripeptidic aldehyde inhibitor Bsc2118 **1** and derivatives with elongated P1 residues **2-4**. b) Visualization of the kinetically activated mechanism as a mouse trap analogy. Hydrophobic interactions between Met⁴⁵ and the ligand's P1 residue lead to conformational changes of the S1 subpocket's backbone. This induced fit strengthens the coordination of a catalytically active water molecule through a network of hydrogen bonds and increases the nucleophilicity of Thr⁴⁰.

To gain detailed insights in the nature of the Met⁴⁵ flexibility, we performed a set of molecular dynamics simulations. Based on the crystal structure of yeast 20S proteasome in complex with **1**, three bound derivatives with iteratively enlarged P1 residues (**2-4**; Figure 1a) and the unbound state were simulated over a timescale of 50 ns (Figure 2a). As expected, the sidechain of Met⁴⁵ shows a high flexibility in absence of any ligand with an RMSD in the range of two to six angstroms. Surprisingly, the simulation with the methyl derivative **2** exclusively revealed a switch between two stabilized states. According to our hypothesis these two stabilized conformations of Met⁴⁵ were interpreted as an unactivated or activated kinetic acceleration, respectively, the latter being highly favored with P1 residues larger than a methyl group. This finding encouraged us to conduct a ligand-based approach to support the calculated data. The diverse numbers of stabilized states due to different P1 lengths were supposed to reflect in the inhibitory activity of a set of ligands. The in vitro testing of derivatives of proteasome inhibitor **1** with iteratively enlarged P1 residues (**2-4**; Figure 1a) allowed a correlation between the size of a residue in the S1 pocket and the resulting inhibitory activity. The length of the respective P1 residue has a direct influence on the dihedral angle α of Met⁴⁵, as shown in Scheme 1 b.

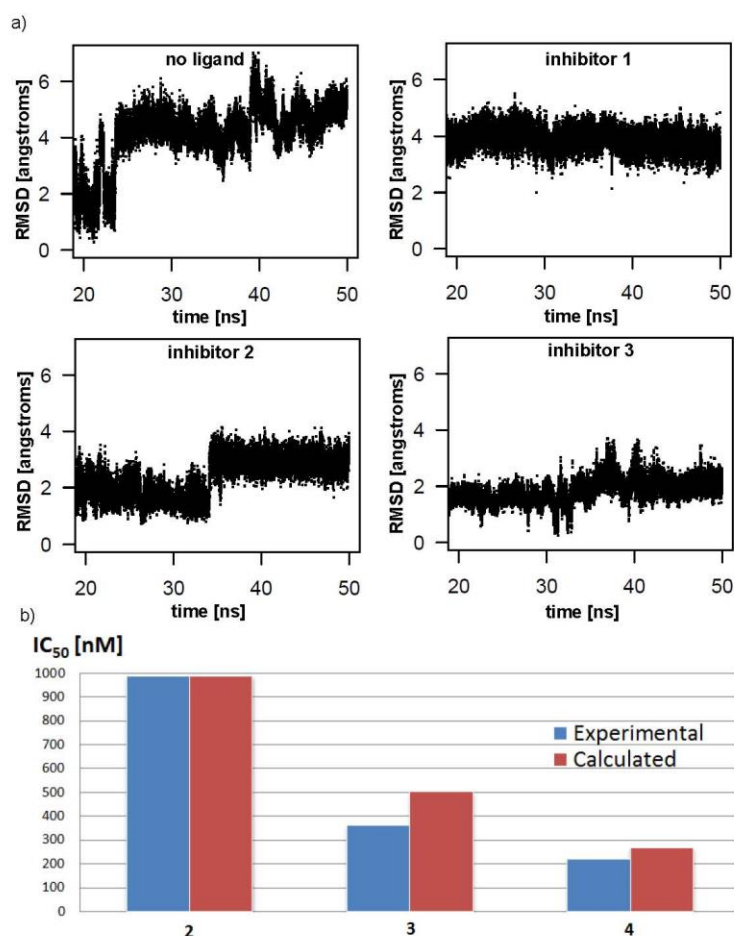


Figure 2: a) Molecular dynamics simulation of the $\beta 5$ subunit with bound inhibitors 1-3 and without any ligand. The RMSD of the Met⁴⁶ sidechain atoms relative to the starting frame is plotted over time. Calculations were done in VMD 1.9.1 and results were plotted in R.^[11] b) Comparison of experimental and calculated IC_{50} values of the inhibitors 2-4. Each experiment was performed as triplicates in two independent measurements ($n=6$). For the in silico calculation the lowest energy conformations were obtained on the basis of the crystal structure of 1 in complex with yeast 20S proteasome. The P1 residue of each inhibitor was adapted and energy minimized to a gradient of 0.01 kcal/mol. The scores (scoring function: London dG) represent relative free energies of ligand binding (ΔG_{rel}) and were converted to relative equilibrium constants by $K_{rel} = \exp(-\Delta G_{rel}/RT)$. Inverse K_{rel} were normalized to fit the dimension of experimental IC_{50} values. See supporting informations for details.

The experimental determination of $\beta 5$ -specific IC_{50} values revealed an erratic increase of the inhibitory activity from P1 = methyl to P1 = ethyl (Figure 2b, experimental). To exclude that this finding had a pure thermodynamic origin due to hydrophobic interactions of the growing P1 residue with the active site, we calculated relative Gibb's free energies of binding in silico (Figure 2b, calculated). An antiproportionality of calculated IC_{50} values with the size of the inhibitor's P1 residue was observed, which is consistent with the assumption that additional methylene units contribute equally to the free energy of a ligand binding process.

In conclusion, the erratic decrease of IC_{50} values coming along with an enlargement of the respective P1 residue from methyl to ethyl confirmed the molecular dynamics simulation. This was our first experimental allusion to the proposed kinetic activation, since it could not be explained by thermodynamic considerations of affinity only.

The next logical step was to switch from inhibition to substrate cleavage for a detailed kinetic study. Therefore, we used a substrate-based approach to identify the minimal P1 length required to induce the kinetic activation. We synthesized fluorogenic 7-amino-4-methylcoumarin (AMC) substrates (**5-7**; Figure 1c), based on the original structure of **1** with iteratively elongated P1 residues. Experimental determination of the Michaelis constant K_M , the turnover rate k_{cat} , the maximum reaction velocity V_{max} and the specificity constant (k_{cat}/K_M) should quantify the impact of a substrate-induced widening of the dihedral angle α of Met⁴⁵ on the $\beta 5$ -activity. To determine the kinetic properties of **5-7** and the widely used reference substrate for the ChT-like activity Suc-LLVY-AMC **S3**, we fluorometrically determined the respective initial velocities of the substrate cleavage in dependency of the substrate concentration and evaluated the data via the well-established Lineweaver-Burke plot (Figure 3). K_M can be interpreted as the dissociation constant of the enzyme/substrate complex, if the formation and dissociation step is much faster than product forming, and thus measures the affinity of a substrate to the active site.^[12] For all four substrates the observed K_M values were found in an almost similar range from 40 μ M for the propyl substrate **7** to 61 μ M for the ethyl substrate **6**. This implies that all substrates, although they have a P1 length distribution from methyl over propyl to tyrosyl, share a similar affinity for the active site without any observable correlation of the P1 length and K_M which is attributed to the flexibility of the Met⁴⁵ sidechain. This flexibility may be able to shape a perfect S1 binding pocket for a variety of hydrophobic P1 residues. According to our initial hypothesis, the kinetic activation due to hydrophobic interactions of a substrate and the Met⁴⁵ sidechain, the turnover rate (k_{cat}) and maximum reaction velocity (V_{max}) should increase with a growing P1 residue. A surprising result of the kinetic analysis was the difference in V_{max} and k_{cat} which was observed as the exact opposite of the expected trend according to the mouse trap hypothesis. Substrates **5** and **6** with small P1 residues showed a twofold higher V_{max} and k_{cat} than the substrates **7** and suc-LLVY-AMC **S3** with spacious P1 sidechains. This opposite trend is also clearly demonstrated by the decrease of the specificity constant k_{cat}/K_M .

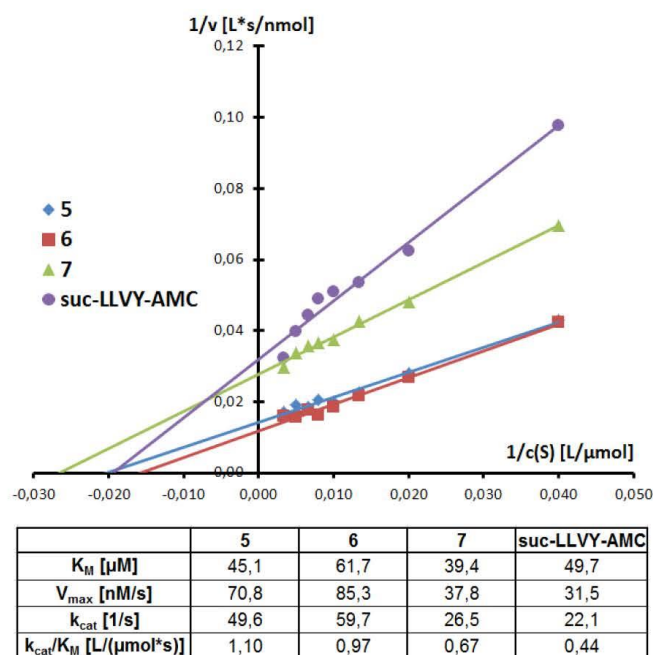


Figure 3: Double reciprocal Lineweaver-Burke plot of substrates **5-7** and suc-LLVY-AMC **S3**. Each experiment was performed as quintuples in two independent measurements ($n=10$). See supplementary materials for experimental and calculation details.

This widely used robust descriptor identifies the best substrate for a given enzyme, in other words, it ranks the specificity of an enzyme for a set of substrates.^[13] This finding is in sharp contrast to the experimentally determined IC_{50} values of respective aldehyde inhibitors **2-4** (Figure 2b, experimental), the previous knowledge derived from countless inhibitor analyses and to the initial mouse trap hypothesis. Substrates with spacious hydrophobic or aromatic P1 residues were thought to be much better substrates than the methyl substrate **5**. The actually observed “small neutral amino acid preferring” (SNAAP) or elastase-like (E-like) activity of the 20S proteasome was first reported by Orlowski et al. in 1993 but at that time referred to an additional and unidentified active subunit.^[14] Despite a report that the $\beta 7$ subunit might be responsible for the SNAAP activity, which was not experimentally proven and suggested from structural data of the bovine 20S proteasome only, it was shown by mutational studies that any observed substrate cleavage is realized by the $\beta 1$, $\beta 2$ or $\beta 5$ subunits.^[15] In these studies, the identification of the SNAAP-responsible active site was not definite since it seemed to depend on the substrate sequence around a potential cleavage site and thus was referred to all three active subunits. The very recently published X-Ray structure of the murine immunoproteasome permits a new explanation why the $\beta 5$ subunit is responsible for this E-like activity. In immunoproteasomes the $\beta 5$ subunit ($\beta 5i$) shows a structural different S1 pocket

compared to its constitutive counterpart ($\beta 5c$).^[16] Whereas in $\beta 5c$ the Met⁴⁵ sidechain is flexible and pointing into the substrate binding channel, it is oriented towards Gln⁵³ due to strong van der Waals interactions in $\beta 5i$. In $\beta 5c$ this Gln⁵³ residue is replaced by serine, which cannot stabilize this orientation of Met⁴⁵ and thus leads to a much smaller S1 binding pocket. Taking all these preliminary studies and the experimental data obtained in this work into account, we conclude that the $\beta 5$ subunit of the constitutive 20S proteasome exerts a SNAAP or E-like rather than a ChT-like activity and should be reclassified. Because of the Met⁴⁵ flexibility substrates with spacious P1 residues are able to reach the active site but their hydrolysis is disfavored compared to substrates bearing a small hydrophobic sidechain on this position. This discrimination maybe due to major structural changes of the protein backbone that are necessary to enlarge the S1 pocket.^[17] Substrate hydrolysis requires fast and dynamic structural changes, as the initial state has to be restored after every cleavage event. This induced fit may be uncritical or even beneficial for inhibiting the $\beta 5$ active site in an irreversible or slowly reversible manner but hinders substrate cleavage. This explains the experimental finding that inhibitors with iteratively enlarged P1 residues show increased inhibitory activities whereas the hydrolytic cleavage of respective substrates with identical peptidic backbones diminishes.

Additionally the examination of the Lineweaver-Burke plots (Figure 3) also brought our attention to a slight deviation from a strict linear dependency for the propyl substrate **7** and more intense for the reference substrate suc-LLVY-AMC **S3**. One feasible explanation for this anomaly is the participation of at least one additional active site to contribute to the substrate cleavage, each showing different values for K_M , k_{cat} or V_{max} . As yeast cell lines lacking the $\beta 1$ and $\beta 2$ active sites of the 20S proteasome were not viable, we had to address the presumed “leakage” for **7** and suc-LLVY-AMC **S3** by other means.

In a first experiment, aldehyde inhibitors (**2-4**; Figure 1a) showed no significant inhibitory activity against any other active site than $\beta 5$ (see supplement). In a subsequent analysis 20S proteasome was preincubated with a high concentration (500 nM) of inhibitor **1** to selectively silence the $\beta 5$ subunit. The hydrolytic activity dropped dramatically for **5** (0.5%) and **6** (0.4%), indicating that these substrates are exclusively cleaved by the $\beta 5$ subunit. Substrate **7** showed a slightly higher remaining hydrolytic cleavage after the $\beta 5$ knockout of 1.2% which may be the reason for the observed deviation from a strict linear Lineweaver-Burke plot. Surprisingly the reference substrate suc-LLVY-AMC **S3**, which is widely used in the majority of ChT-like activity assays, showed a remaining activity of 12% under the same conditions. This is consistent with the obvious non-linear behaviour of the Lineweaver-Burke plot. Unselective cleavage of the established reference substrate may be critical when

suc-LLVY-AMC **S3** is used to rank the selectivity and potency of inhibitors for the ChT-like activity. For this reason substrate **5**, Z-Leu-Asp(OtBu)-Ala-AMC, is preferable for these applications and the substitution of suc-LLVY-AMC **S3** is recommended.

Literaturverzeichnis:

- [1] A. Hershko, E. Leshinsky, D. Ganoth, H. Heller, *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 1619-1623.
- [2] L. Borissenko, M. Groll, *Chem. Rev.* **2007**, *107*, 687-717.
- [3] a) J. Driscoll, M. G. Brown, D. Finley, J. J. Monaco, *Nature* **1993**, *365*, 262-264; b) P. M. Voorhees, E. C. Dees, B. O'Neil, R. Z. Orłowski, *Clin. Cancer Res.* **2003**, *9*, 6316-6325.
- [4] a) K. M. Kortuem, A. K. Stewart, *Blood* **2013**, *121*, 893-897; b) R. C. Kane, A. T. Farrell, R. Sridhara, R. Pazdur, *Clin. Cancer Res.* **2006**, *12*, 2955-2960.
- [5] S. Meiners, A. Ludwig, V. Stangl, K. Stangl, *Med. Res. Rev.* **2008**, *28*, 309-327.
- [6] H. A. Braun, S. Umbreen, M. Groll, U. Kuckelkorn, I. Młynarczuk, M. E. Wigand, I. Drung, P.-M. Kloetzel, B. Schmidt, *J. Biol. Chem.* **2005**, *280*, 28394-28401.
- [7] J. Sterz, C. Jakob, U. Kuckelkorn, U. Heider, M. Mieth, L. Kleeberg, M. Kaiser, P.-M. Kloetzel, O. Sezer, I. von Metzler, *Eur. J. Haematol.* **2010**, *85*, 99-107.
- [8] E. J. A. M. Sijts, P. M. Kloetzel, *Cell. Mol. Life Sci.* **2011**, *68*, 1491-1502.
- [9] a) F. Luciani, C. Kesmir, M. Mishto, M. Or-Guil, R. J. De Boer, *Biophys. J.* **2005**, *88*, 2422-2432; b) M. Mishto, F. Luciani, H.-G. Holzhütter, E. Bellavista, A. Santoro, K. Textoris-Taube, C. Franceschi, P. M. Kloetzel, A. Zaikin, *J. Mol. Biol.* **2008**, *377*, 1607-1617; c) A. K. Nussbaum, C. Kuttler, K.-P. Hader, H.-G. Rammensee, H. Schild, *Immunogenetics* **2001**, *53*, 87-94; d) C. Kesmir, A. K. Nussbaum, H. Schild, V. Detours, S. Brunak, *Protein Eng.* **2002**, *15*, 287-296.
- [10] M. Groll, C. R. Berkens, H. L. Ploegh, H. Ovaa, *Structure* **2006**, *14*, 451-456.
- [11] a) *R Core Team. R: A language and environment for statistical computing.*, R Foundation for Statistical Computing, Vienna, Austria., **2013**; b) W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33-38.
- [12] W. W. Chen, M. Niepel, P. K. Sorger, *Genes Dev.* **2010**, *24*, 1861-1875.
- [13] R. Eisinger, M. J. Danson, D. W. Hough, *Trends Biotechnol.* **2007**, *25*, 247-249.
- [14] M. Orłowski, C. Cardozo, C. Michaud, *Biochemistry* **1993**, *32*, 1563-1572.
- [15] a) M. Unno, T. Mizushima, Y. Morimoto, Y. Tomisugi, K. Tanaka, N. Yasuoka, T. Tsukihara, *Structure* **2002**, *10*, 609-618; b) T. P. Dick, A. K. Nussbaum, M. Deeg, W. Heinemeyer, M. Groll, M. Schirle, W. Keilholz, S. Stevanović, D. H. Wolf, R. Huber, H.-G. Rammensee, H. Schild, *J. Biol. Chem.* **1998**, *273*, 25637-25646.
- [16] Eva M. Huber, M. Basler, R. Schwab, W. Heinemeyer, Christopher J. Kirk, M. Groettrup, M. Groll, *Cell* **2012**, *148*, 727-738.
- [17] E. M. Huber, M. Groll, *Angew. Chem. Int. Ed.* **2012**, *51*, 8708-8720.

3.3 Entwicklung von DOCKTITE, einer Methode zum automatisierten Docking kovalenter Liganden

Der Inhalt dieses Kapitels wurde bereits veröffentlicht:

Christoph Scholz, Sabine Knorr, Kay Hamacher, Boris Schmidt, "DOCKTITE – A Highly Versatile Step-by-Step Workflow for Covalent Docking and Virtual Screening in the Molecular Operating Environment", *Journal of Chemical Information and Modeling* **2015**, 55, 398-406.

Mit freundlicher Genehmigung der American Chemical Society.

Das Protein/Ligand-Docking gehört zu den wichtigsten und meistgenutzten Methoden des strukturbasierten *Molecular Modeling*. Dennoch ist die Auswahl an Softwarelösungen für das kovalente Docking sehr überschaubar, die Methoden oft wenig benutzerfreundlich und nur auf wenige Spezialfälle anwendbar. Die Motivation zur Entwicklung von DOCKTITE war es, genau diese Lücke zu schließen. DOCKTITE sollte keinerlei Restriktionen bezüglich elektrophiler Kopfgruppen aufweisen, neue Bindungsmechanismen sollten benutzerfreundlich hinzugefügt werden können, es sollten schnelle und akkurate Ergebnisse erzeugt und virtuelle Screenings größerer Bibliotheken möglich gemacht werden. Darüber hinaus sollte DOCKTITE frei verfügbar und *open source* sein, damit fortgeschrittene Nutzer eigene maßgeschneiderte Änderungen einfügen können.

Die Softwareumgebung MOE ist aus mehreren Gründen als Plattform für DOCKTITE besonders geeignet: (1) Die Quellcodes ihrer Anwendungen sind *open source*, (2) durch die graphische Oberfläche ist eine hohe Benutzerfreundlichkeit gegeben, (3) durch die modulare Natur von MOE können bereits implementierte Programme wie die *MedChem-Transformations*, der Pharmakophor-Editor und der Dockingalgorithmus problemlos genutzt werden und (4) MOE baut auf einer eigens für wissenschaftliche Ansprüche entwickelten Programmiersprache – der *Scientific Vector Language* (SVL) – auf, die relativ leicht zu erlernen ist.

DOCKTITE ist ebenfalls modular aufgebaut. Im ersten Schritt, dem *Warhead Screening*, wird eine Ligandendatenbank vollautomatisch nach Verbindungen mit elektrophiler Kopfgruppe durchsucht, die Kopfgruppe in die gebundene Geometrie überführt und das elektrophile Reaktionszentrum für den nächsten Schritt markiert. Hierfür wird

eine mittels SVL für Datenbanken angepasste Version der MOE-internen *MedChem-Transformations* genutzt – ein Unterprogramm, mit dem sich anhand vorgegebener Transformationsregeln die Struktur einer Verbindung verändern lässt. Das Besondere ist, dass diese Transformationsregeln in einem 2D-Zeichenprogramm für chemische Strukturen (ChemDraw, AccelrysDraw, etc.) erstellt werden können und man keine unübersichtlichen zeilenbasierten Notationen wie SMARTS oder SMIRKS benötigt.^[116]

Der zweite Schritt stellt den Schlüsselschritt von DOCKTITE dar. Zunächst wird die nukleophile Aminosäure des Rezeptors (bsp. Serin, Cystein oder Threonin) vom Rest des Proteins getrennt und kovalent mit dem Liganden verbunden. Für jedes Schweratom der Aminosäure werden gleichzeitig Pharmakophordescriptoren erstellt, welche im anschließenden Dockingschritt die nukleophile Aminosäure fixieren. Das hierbei vollautomatisch erstellte Ph4-Modell kann nach Belieben verändert oder erweitert werden, je nach aktuellem Kenntnisstand bezüglich essentieller Wechselwirkungen des Liganden mit dem Zielprotein (siehe Abschnitt 1.2.1). Mittels eines pharmakophorbasierten *Placement*-Algorithmus werden anschließend die chimären Verbindungen, bestehend aus Ligand und nukleophiler Aminosäure, in die Bindungstasche des Rezeptorproteins gedockt (siehe Abschnitt 1.2.3). Der Dockingschritt beinhaltet eine Energieminimierung der Bindungsgeometrien, die entweder kraftfeld- oder GRID-basiert durchgeführt werden kann – je nach vorhandener Zeit und Rechenleistung. Die Identifizierung des plausibelsten Bindungsmodus und die Berechnung der zugehörigen freien Bindungsenthalpie werden über einen innovativen *Multi-Scoring*-Ansatz realisiert. Da wissensbasierte *Scoring*-Funktionen für die Identifizierung einer realitätsnahen Bindungsgeometrie im Vorteil sind, wird hierfür zunächst der *Drug Score Extended* (DSX) eingesetzt.^[49] Die Berechnung der resultierenden freien Bindungsenthalpie wird, nach einer vollautomatischen Trennung von Ligand und nukleophiler Aminosäure, über die MOE-internen empirischen *Scoring*-Funktionen *London dG* und *Affinity dG* durchgeführt.

Die Validierung der Prädiktivität realitätsnaher Bindungsmodi erfolgte zunächst mittels eines Datensatzes, der auch von zahlreichen anderen Dockingprogrammen genutzt wurde (CovDock, DOCKKovalent, CovalentDock, Autodock und GOLD) und somit einen direkten Vergleich ermöglichte. DOCKTITE wurde hierbei lediglich von CovDock leicht übertroffen (Tabelle 1).

Tabelle 1: Vergleich verschiedener Dockingprogramme mittels eines Referenzdatensatzes, bestehend aus 76 experimentell bestimmten Protein/Ligand-Komplexen.

	DOCKTITE	CovDock	DOCKkovalent	CovalentDock	Autodock	GOLD
RMSD^a	2.4 Å	1.8 Å	2.9 Å	3.4 Å	3.5 Å	4.0 Å
< 2 Å^b	48.7%	63.2%	37.7%	-	-	-

^aMittlere RMSD-Werte der vorhergesagten Bindungsgeometrien, bezogen auf experimentell bestimmte Protein/Ligand-Komplexe. ^bAnteil an vorhergesagten Bindungsgeometrien mit einer RMSD von weniger als 2 Å.

Wie bereits in Abschnitt 1.2.3 erläutert, ist CovDock aufgrund der Grundhypothese, dass sich der gebundene Bindungsmodus nicht sehr vom ungebundenen unterscheidet, auf viele Wirkstoffklassen nicht anwendbar. Um genau diese Vielseitigkeit von DOCKTITE zu demonstrieren, wurde ein weiterer Datensatz kovalenter Protein/Ligand-Komplexe äußerst erfolgreich simuliert, welcher 35 Liganden mit insgesamt 13 verschiedenen elektrophilen Kopfgruppen und Kinasen sowie Proteasen mit Serin, Threonin und Cystein als nukleophile Aminosäuren beinhaltet.

Die Fähigkeit, in einem dockingbasierten virtuellen Screening aktive von inaktiven Verbindungen zu unterscheiden, wurde anhand eines Datensatzes, bestehend aus 42 inaktiven und 7 aktiven Inhibitoren des 20S-Proteasoms, validiert. Die Auswertung erfolgte über eine sog. *Receiver Operating Characteristic* (ROC)-Kurve, deren Fläche unter der Kurve ein Maß für die Effizienz eines virtuellen Screenings ist. DOCKTITE erreichte hierbei einen ausgezeichneten Wert von 0,8 (1 = ideal und 0,5 = zufällig).

Die Validierung der Affinitätsbestimmung wurde durch ein *Ranking* zehn chemisch verwandter Inhibitoren der c-Src-Kinase realisiert. Hierbei hängt der Erfolg dramatisch von den verwendeten *Scoring*-Funktionen ab und ist aufgrund der zugrunde liegenden Vereinfachungen der physikochemischen Abläufe extrem fehleranfällig. Dennoch konnte eine signifikante Korrelation zwischen experimentell bestimmten apparenten Bindungsaffinitäten (pIC₅₀) und dem DOCKTITE-Score erreicht werden ($R^2 = 0,649$ / $p < 0,005$).

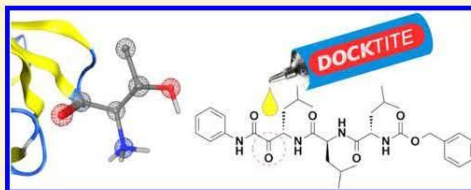
Beitrag C. Scholz: Idee, Verfassen des Manuskripts, Programmierung der SVL-Skripte, Implementierung von 23 Bindungsmechanismen, Validierung der Methode.

DOCKTITE—A Highly Versatile Step-by-Step Workflow for Covalent Docking and Virtual Screening in the Molecular Operating Environment

Christoph Scholz,^{*,†} Sabine Knorr,^{||} Kay Hamacher,^{||} and Boris Schmidt^{*,†}[†]Clemens Schöpf-Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Alarich-Weiss-Strasse 4, 64287 Darmstadt, Germany^{||}Computational Biology & Simulation, Technische Universität Darmstadt, Schnittspahnstraße 10, 64287 Darmstadt, Germany

Supporting Information

ABSTRACT: The formation of a covalent bond with the target is essential for a number of successful drugs, yet tools for covalent docking without significant restrictions regarding warhead or receptor classes are rare and limited in use. In this work we present DOCKTITE, a highly versatile workflow for covalent docking in the Molecular Operating Environment (MOE) combining automated warhead screening, nucleophilic side chain attachment, pharmacophore-based docking, and a novel consensus scoring approach. The comprehensive validation study includes pose predictions of 35 protein/ligand complexes which resulted in a mean RMSD of 1.74 Å and a prediction rate of 71.4% with an RMSD below 2 Å, a virtual screening with an area under the curve (AUC) for the receiver operating characteristics (ROC) of 0.81, and a significant correlation between predicted and experimental binding affinities ($\rho = 0.806$, $R^2 = 0.649$, $p < 0.005$).



INTRODUCTION

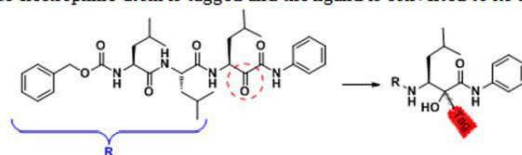
Computational tools for protein/ligand docking have been developed predominantly for noncovalent modulators of a specific target protein.^{1,2} However, the formation of a covalent bond is essential for a number of successful drugs, such as the first and second generation proteasome inhibitors bortezomib and carfilzomib, which are FDA-approved drugs for the treatment of multiple myeloma and mantle cell lymphoma.³ Improved biochemical efficiency, selectivity, lower and less frequent dosing during patient treatment and improved pharmacodynamics are some of the potential benefits of covalent target inhibition.⁴ In the scope of proteasome research, we were looking for suitable *in silico* methods to identify novel covalent inhibitors of this important threonine protease. Even though a number of popular docking programs employ covalent docking implementations, the structural versatility, accuracy, and automation needed for a large-scale docking-based virtual screening are rare. GOLD⁵ for example uses a linker atom which is part of the ligand and the receptor. Schröder et al. expanded this method for virtual screenings by establishing an automated ligand modification script utilizing the software package CACTVS.⁶ AutoDock⁷ enables the user to choose between two methods for covalent docking. The grid-based method uses an additional grid map to fix the linker atom of the ligand, while the alternative method considers the ligand as a flexible side chain of the receptor. To our knowledge, an automated ligand preparation method has not been provided so far. The first systematic comparison of different covalent

docking methods was reported by Ouyang et al.⁸ Their method makes use of a dummy atom to ensure the covalent bond and explicitly scores its energy contribution using a Morse potential, which was parametrized for the two examined binding reactions. The most recent methods were published by Zhu et al. (CovDock) and London et al. (DOCKovalent).^{9,10} In DOCKovalent, electrophilic warheads are identified by SMARTS patterns and the respective ligands are converted to their bound state prior to the covalent docking. The overall ligand preparation methodology makes use of five different programs, which might be discouraging to new users. To address a wide range of users, the authors provide an easier to use web-accessible version of their method, where predefined ligand libraries can be docked into individual receptor structures.¹¹ The docking itself is conducted using the software DOCK3.6 and a physics-based scoring function.¹² Zhu et al. use different tools of the software suite Schrödinger to mimic distinct stages of covalent inhibitor binding. The first step is a classical noncovalent docking with an alanine mutation of the nucleophilic side chain followed by an automated bond formation and a second docking step with the covalent bond in place. Their working hypothesis is that a covalently bound ligand has to adopt an energetically favorable unbound pose before bond formation occurs and that these unbound poses do not change dramatically during the reaction pathway, because

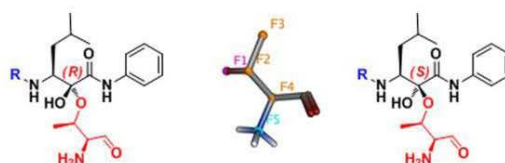
Received: November 12, 2014

Published: December 26, 2014

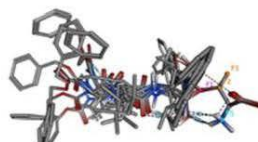
Step 1: Warhead Screening. A database of ligands is searched for electrophilic warheads. Subsequently, the electrophilic atom is tagged and the ligand is converted to its bound shape:



Step 2: Side chain Attachment. The tagged ligand atom is now attached to the nucleophilic side chain of the receptor and, for prochiral warheads, stereoisomers of the chimeric molecules are generated. Additionally DOCKTITE generates an automatic pharmacophore (ph4) model for the attached residue and analyzes the active site automatically. This model is used for an exact positioning during the ph4-guided docking process.



Step 3: Pharmacophore-guided Docking. The ph4 placement method is responsible for an exact positioning of the nucleophilic side chain during the docking step.



Step 4: Side chain Cleavage & Pose Rescoring. Estimating the final docking scores leads to best results with free ligands, disconnected from their attached side chain. After the fully automated cleavage step, the rescoring is realized by a newly developed consensus scoring approach using MOE-internal and the external knowledge-based scoring function DSX.

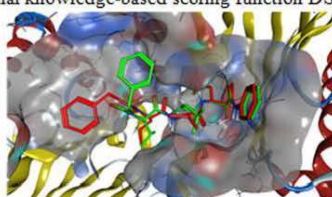


Figure 1. Overview of the DOCKTITE workflow.

conformational sampling is done solely prior the noncovalent docking step. This method works well in most cases but has predictable problems with ligands that change their conformation as a result of target binding, for example due to a distortion of conjugated systems, ring openings, or elimination of leaving groups, leading to limited warhead versatility of CovDock.

In this work we present DOCKTITE, the implementation of a highly versatile covalent docking workflow, capable of virtually screen large databases of potential lead structures, within the Molecular Operating Environment (MOE) via SVL-scripts.¹³ The unique step-by-step method allows the user to

keep control over each step of a docking project and to tailor it individually. The process is guided by an easy-to-use graphical user interface, so DOCKTITE addresses beginners and advanced users equally. The warhead screening step includes browsing a ligand database for user-defined covalent warheads which are afterward tagged and converted into a bound state. After connecting the potential ligands to the covalently bound residue of the receptor, taking additionally generated stereoisomers into account, a pharmacophore-based docking with optional force field refinement is conducted. Subsequently, the chimeric poses are cleaved and rescored by an innovative consensus scoring approach using MOE-internal empirical

scoring functions and the external knowledge-based scoring function DSX.¹⁴ To demonstrate the versatility and accuracy of DOCKTITE, we extensively challenged it to identify near-native binding modes of a large set of ligands with various electrophilic warheads, including aldehydes, boronates, vinyl sulfones, vinyl amides, α -halo ketones, α -keto amides, β -lactones, phosphonofluoridates, benzoxazinones or isocoumarins, and on a variety of proteases and kinases bearing serine, cysteine, or threonine residues as nucleophilic attachment points. In addition to the validation of docking power, we demonstrate that DOCKTITE is an extremely fast and useful tool for large-scale virtual screening purposes, able to differentiate binders from nonbinders and to rank active compounds regarding their experimentally determined binding affinity values in a congeneric series of ligands.

MATERIALS AND METHODS

Workflow. A fast and user-friendly method for covalent docking without restrictions upon warhead or receptor classes is presented, being able to deal with large-scale virtual screening tasks. Furthermore, advanced users may modify our open source program to suit their demands. MOE is predestined for this purpose, as the source codes of its applications are natively open source and written in the rather easy to learn *Scientific Vector Language* (SVL) that was exclusively developed for MOE.¹⁵ The overall workflow is shown in Figure 1.

Step 1: Warhead Screening. In the first step, a database of ligands is screened for predefined electrophilic warheads. Subsequently, the electrophilic atom is tagged and the ligand is converted to its bound shape. All ligand manipulations are done fully automatic by a modified version of MOE's *MedChem Transformations* (MCT), an application initially used to discover novel chemical structures by applying a set of transformation rules to existing ligands. A major advantage of this method is the intuitive way to customize binding reactions. Even for simple warheads, the use of line notations like SMIRKS can be very confusing for the user while in DOCKTITE the reaction is simply drawn in ChemDraw or freeware 2D-Drawers that support the standard reaction format RXN.^{16,17} SVL-based modifications were necessary to use a database of ligands instead of a single ligand as the input, whereas the overall algorithm of MCT was left untouched. It consists of a substructure search to match the *query* (left-hand side) of each transformation rule to the input ligands and a replacement of the matched atoms by those of the *result* (right-hand side), while each individual transformation rule indicates whether query atoms or bonds are copied rather than replaced. The tagging is realized by the introduction of a unique atom that has to be absent in all input structures and mimics the nucleophilic atom of the covalent receptor attachment point (see Scheme 1). Tantalum (Ta) is used for temporary tagging, as this element is usually not present in drug-like ligand sets. Furthermore, two additional tagging elements are used in the same manner: yttrium (Y) to define leaving groups and germanium (Ge) to mimic tetravalent boron atoms which are natively not allowed in MCT. Even though all three tagging elements are very uncommon in drug development, they can easily be substituted by the user in the unlikely case of interfering with the ligand set.

Up to now we have implemented 21 of the most common warhead classes known for covalent inhibitors in DOCKTITE, including binding mechanisms featuring rearrangements or leaving groups.^{3,18} It is important to note that this set is easily

Scheme 1. Tagging Elements Used in DOCKTITE^a



^a(a) Tagging elements are placed in step 1: germanium (Ge) to mimic a tetravalent boron (B) atom (purple), tantalum (Ta) to tag the electrophilic atom of the warhead (red), and yttrium (Y) to separate a leaving group from the rest of the molecule (blue). (b) Tagging elements serve as signals in step 2 and are replaced: germanium is replaced by a tetravalent boron atom (purple), tantalum is replaced by the nucleophilic side chain of the receptor (red), and the leaving group is deleted (blue).

expandable on demand (see Figure 2 for warheads used in this work and the Supporting Information for a complete list of implemented warheads).

Step 2: Side Chain Attachment. This is the key step of DOCKTITE, as it includes a variety of modifications and preparations for the final covalent docking. The residue bearing the nucleophilic side chain of the prepared receptor is disconnected from the rest of the protein and fixed in its native position by pH4 constraints. For each individual heavy atom of the residue a pH4 feature is created with a standard size of 0.4 Å. Concurrently the active site is identified and analyzed by means of size and shape to generate an additional pharmacophore feature, requiring any ligand heavy atom, 1.5 Å away from the nucleophilic attachment point projected in the direction of the active site to guide the later pharmacophore-based docking (see Figure 3A). An important advantage of DOCKTITE is that additional information on essential interactions or other known constraints can be easily implemented by user-defined pharmacophore features in this step (see Figure 3B).

The tags, applied in step 1, serve as signals for ligand modifications. After deletion of a possible leaving group and converting a possible Ge-tag to boron, the nucleophilic residue is connected to the Ta-tagged electrophilic attachment point of the ligand. Then, the resulting chimeric molecule is energy-minimized using the force field loaded in MOE. If the warhead of the ligand is prochiral, stereoisomers are generated and treated as individual ligands in all following steps.

Step 3: Pharmacophore-Guided Docking. The final docking step is conducted using MOE's *DOCK* application which can be divided into the following subroutines: (1) generation of conformations, (2) placement, (3) scoring, (4) pose refinement, and (5) rescoring. Conformations are generated by applying a collection of preferred torsion angles to rotatable bonds. A systematic search is carried out covering all combinations of angles on a grid, if this will result in less than 5000 conformers. Otherwise a stochastic sampling of conformations is conducted. The pharmacophore placement method is responsible for an exact positioning of the nucleophilic residue and guides the docking and refinement step. For pharmacophore queries with high complexity, like those obtained in DOCKTITE, the pharmacophore search engine of MOE is used to orient the ligand conformations in the active site. The generated poses are ranked with the empirical scoring function *Affinity dG* which estimates enthalpic contributions to the free energy of binding by a linear function consisting of terms taking hydrogen bonds, ionic interactions, metal ligations, hydrophobic interactions, and

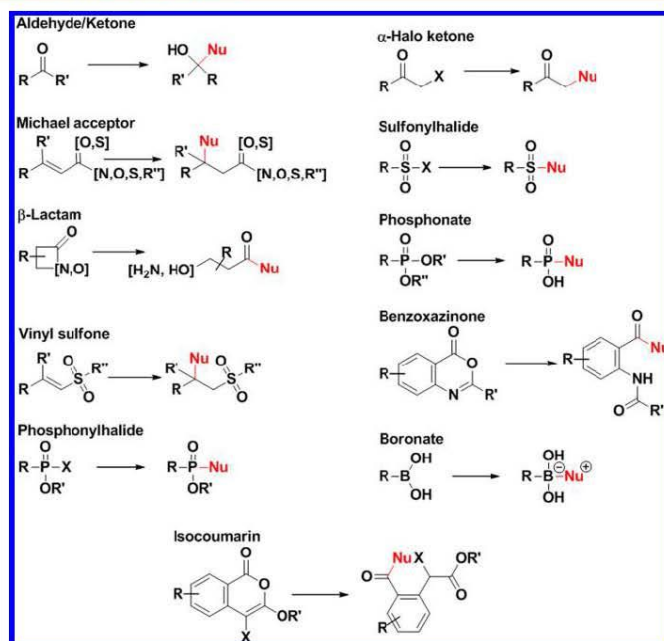


Figure 2. Selection of electrophilic warheads implemented in DOCKTITE (see the Supporting Information for a complete list). Nu: nucleophilic attachment point of the receptor. R, R', R'': any substituent with terminal carbon. X: any halogene. Note: no stereoisomers and not all possible substitutions are shown for reasons of clarity. The covalent bond is shown in red.

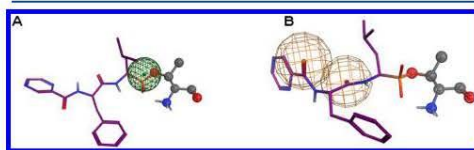


Figure 3. Pharmacophore models mapped on proteasome inhibitor bortezomib (PDB ID: 2fl6): A automatic ph4 model used in pose prediction and B extended ph4 model with additional constraints used in virtual screening. Purple: ligand carbon atoms. Gray: nucleophilic residue carbon atoms. Ph4 features are represented as colored spheres: oxygen (red), nitrogen (blue), carbon (gray), any heavy atom (green), H-bond acceptor (orange).

unfavorable polar/hydrophobic contacts into account. The 100 best scored poses are further refined inside the active site either by energy minimization (in pose prediction, exhaustive virtual screening, and ranking of congeneric ligands) or by grid minimization (in fast virtual screening). For the refinement by energy minimization the AMBER12:EHT force field is used, which is parametrized for both proteins and small molecules.^{19,20} Using force field refinement, a docking process of one ligand approximately takes 0.5–1 CPU hours on a medium-priced workstation to complete. For typical docking-based virtual screenings with thousands or even millions of ligands, this time scale is not applicable. For these purposes the poses are refined by grid minimization which reduces the CPU time to 10–20 s per ligand. The refined poses are again scored

using the *Affinity dG* scoring function. It is noteworthy that, due to the modular nature of the DOCKTITE workflow, this step can be conducted in any modeling software being able to do a pharmacophore-based placement.

Step 4: Side Chain Cleavage and Pose Rescoring. The estimation of final docking scores leads to best results with free ligands, disconnected from their attached side chain, because small conformational alterations in the residue part of the chimeric molecule may have a huge impact of calculated binding affinities. Subsequent to the fully automated cleavage step a multiscore approach, called consensus scoring, is used to (1) identify the near-native docking pose and (2) estimate the respective free energy of binding. The first task, which is a prerequisite for the second one, is done by the knowledge-based scoring function DSX. In contrast to their empirical counterparts, e.g. *Affinity dG*, knowledge-based scoring functions are not trained with binding affinities. These functions calculate the total score as a sum of statistical potentials, which are derived from a database of known protein/ligand complexes with solved 3D structures, in this case from the Protein Data Bank (PDB).²¹ As the score of DSX reflects native binding geometries, predicting a near-native conformation is its key skill while it is not designed to calculate binding energies in particular. The second task is performed by the *Affinity dG* or *London dG* scoring function implemented in MOE, considering only the top pose according to DSX. This consensus scoring approach employed in DOCKTITE combines the superior binding mode predictivity of DSX with greater accuracy of energy estimation by empirical scoring

Table 1. Pose Prediction Results for the 76 Complexes and Comparison with Previously Reported Covalent Docking Programs

	DOCKTITE	CovDock	DOCKKovalent ^d	CovalentDock	Autodock	GOLD
Top pose ^a	2.4 Å	1.8 Å	2.9 Å	3.4 Å	3.5 Å	4.0 Å
<2 Å ^b	48.7%	63.2%	37.7%	n.a. ^e	n.a.	n.a.
<1 Å ^c	5.3%	31.6%	6.6%	n.a.	n.a.	n.a.
Best in top 10 ^d	1.7 Å	1.4 Å	n.a.	1.9 Å	2.5 Å	3.4 Å
<2 Å ^b	73.7%	77.6%	n.a.	n.a.	n.a.	n.a.
<1 Å ^c	18.4%	50.0%	n.a.	n.a.	n.a.	n.a.

^aAll heavy atom RMSD values are calculated in angstroms from the unmodified X-ray structure. ^bPercentage of correctly predicted poses, when a threshold of 2 Å is defined. ^cPercentage of correctly predicted poses, when a threshold of 1 Å is defined. ^dTwo β -lactam and 13 Michael acceptor inhibitors were excluded from the validation of DOCKKovalent.¹⁰ ^eData not available.

functions, the latter being rather useless when calculated for a "wrong" docking pose.

Ligand Preparation. In general, all ligands are used in their prereacted form. Manual structure modifications are done with the *MOE Builder* if necessary; otherwise, they are directly used from their source database. Hydrogens are added and charges assigned using the potential setup interface. Low-energy conformations of ligands are prepared via energy minimization using the *MMFF94x* force field and a distance-dependent dielectric for solvation treatment.²²

Receptor Preparation. The X-ray structures are prepared with the *LigX* interface of MOE. *LigX* combines several structure preparation steps including deletion of incidental water molecules, *Protonate 3D*, assigning ASN/GLN/HIS-“Flips”, geometry optimization, and energy minimization (RMSD gradient = 0.1 kcal/mol) of the prepared receptor. The *Protonate 3D* application assigns protonation states from a discrete collection of states by optimizing the titration free energy of all titratable groups in the context of an all-atom model of a macromolecular structure (including ligands and solvent).²³ The Generalized Born/Volume Integral Electrostatics model is used for long-range interactions and solvation effects.²⁴ For energy minimization the *AMBER12:EHT* force field is used.^{19,20}

Pharmacophore Preparation. In case of unknown essential interactions between the ligand and the active site, the automatic ph4 model can be used or extended with an additional *AtomQ* feature (which stands for any heavy atom) placed on a dummy atom that is generated with the *MOE SiteFinder* application. In virtual screening and ranking of congeneric ligands, a more precise ph4 model with additional essential features is used, which represents the typical state of knowledge during a drug development process (see Figures 3B and 6B).

Validation Sets. The validation of DOCKTITE was subdivided into three core disciplines exceptionally important for drug design: prediction of near-native binding modes, fast and accurate virtual screening, and ranking of ligands in a congeneric series regarding their experimentally determined binding affinities. Where possible, a comparison of our method with previously published protocols was conducted using the same data sets. However, a new data set for virtual screening was used and the previously reported validation set of CovDock was discarded.⁹ We partly disagree with the compound selection, as nothing was known about the inhibition potency of the decoy compounds. The prediction of near-native binding modes was tested using two data sets. The first set was published by Ouyang et al. and contains 76 covalent complexes, subdivided in 13 Michael acceptor and 63 β -lactam inhibitors.⁸ The advantage of this data set is that five previously published

methods for covalent docking were validated against it: GOLD, Autodock, CovalentDock, CovDock, and DOCKKovalent.^{8–10} Nevertheless, warhead and receptor variability in this set of compounds is not representative for today's repertoire of covalent inhibitors. Hence, the versatility of DOCKTITE was validated against a broad data set containing 35 complexes, spanning a majority of known covalent warheads used in drug discovery today. The database is partly derived from the work of Zhu et al. and extended by 11 warhead classes. To our knowledge, most of them have never been explored in a systematic covalent docking study before.⁹ To validate the applicability of DOCKTITE in virtual screenings, we used a set of 49 known covalent 20S proteasome inhibitors derived from the public database BindingDB.^{25–27} The search for 20S proteasome inhibitors ($\beta 5$ -subunit) resulted in 63 nonduplicate entries with published K_i or IC_{50} values. This data set was subdivided into 7 active ($K_i/IC_{50} \leq 500$ nM) and 42 inactive compounds ($K_i/IC_{50} \geq 10$ μ M), which are further annotated as actives and inactives, respectively (see the Supporting Information for a complete list).^{28–33} The validation of DOCKTITE's scoring power was conducted by ranking a congeneric series of 10 c-Src kinase inhibitors with an acrylamide (Michael acceptor) warhead according to their experimentally determined apparent binding affinities (IC_{50}).³⁴

RESULTS AND DISCUSSION

Prediction of Native Binding Modes. Reliability in pose prediction is a prerequisite for other tasks like virtual screening or ranking of congeneric ligands and a focus of the DOCKTITE evaluation. The heavy atom RMSD between a predicted pose and the pose observed in the unmodified crystal structure was used as a metric for prediction accuracy. To compare DOCKTITE with previously reported methods for covalent docking, we validated it against the same 76 complexes used by Zhu et al. and Ouyang et al. as well (see Table 1 and the Supporting Information).^{8,9} The pose prediction accuracy (top pose) of DOCKTITE turned out to be superior to those of CovalentDock, AutoDock, and GOLD with an RMSD benefit of at least 1 Å. However, we were unable to outperform the results of CovDock. This special test set consists only of two kinds of covalent ligands (Michael acceptors and β -lactams), and their binding modes seem to follow the assumption made by Zhu et al., that a ligand keeps its conformation after covalent binding, almost perfectly. The main focus during the development of DOCKTITE was its versatility with respect to electrophilic warhead and receptor classes. This versatility was challenged with an additional validation set of 35 known covalent complexes, partly derived from the work of Zhu et al. and expanded with ligands of 11 warhead classes.⁹

Table 2. Redocking of 35 Covalent Complexes, Bearing 14 Distinct Electrophilic Warheads

PDB	input complex			predicted pose			best pose in top 10	
	warhead	residue	R/S ^a	score ^{DSX}	RMSD ^b	R/S ^a	score ^{DSX}	RMSD ^b
2awz	acrylamide	Cys366	R	-65.88	2.45	R	-62.29	1.13
2ax0	acrylamide	Cys366	R	-85.06	1.00	R	-85.06	1.00
2axl	acrylamide	Cys366	R	-65.47	1.68	R	-57.92	1.20
2e14	vinyl ketone	Cys166	S	-98.14	1.76	R	-98.14	1.76
2hwo	acrylamide	Cys345		-87.63	1.94		-77.17	1.37
2hwp	acrylamide	Cys345		-61.69	2.38		-61.69	2.38
2jiv	acrylamide	Cys797	S	-187.30	1.74	R	-183.41	1.59
2qlq	acrylamide	Cys345	R	-99.80	2.73	R	-89.07	1.13
3c9w	vinyl ketone	Cys164	R	-102.06	1.08	R	-99.38	1.07
3ika	acrylamide	Cys797		-23.00	4.46		-23.00	4.46
3lok	acrylamide	Cys345		-75.21	1.60		-69.29	0.78
3oyp	acrylamide	Cys159		-99.05	2.38		-88.98	2.35
3t9t	acrylamide	Cys442	S	-132.14	1.39	R	-128.62	0.65
2grf	α -keto amide	Ser139	S	-188.85	1.49	R	-165.02	1.33
2obo	α -keto amide	Ser139	S	-121.07	0.69	S	-117.63	0.68
2oc1	α -keto amide	Ser139	S	-142.84	1.47	R	-142.84	1.47
2p59	α -keto amide	Ser1165	S	-146.83	1.36	S	-142.90	1.25
3lox	α -keto amide	Ser139	S	-104.97	2.17	S	-102.97	1.31
2oc7	α -keto amide	Ser139	S	-104.63	1.26	S	-89.36	1.05
2oc8	α -keto amide	Ser139	S	-118.59	1.23	S	-118.59	1.23
2oc0	α -keto amide	Ser139	S	-121.24	0.91	S	-116.24	0.83
3knx	α -keto amide	Ser139	S	-166.90	0.99	S	-158.63	0.96
1w3c	α -keto amide	Ser139	S	-85.03	2.15	R	-83.74	1.45
2f9u	α -keto amide	Ser139	S	-154.44	3.43	S	-154.44	3.43
1zcm	α -haloketone	Cys115		-70.53	4.00		-47.26	1.55
2g8e	aldehyde	Cys115	R	-88.81	0.93	R	-88.81	0.93
2fl6	boronate	Thr1		-123.52	1.24		-123.52	1.24
4lqi	β -lactone	Thr1		-58.49	1.61		-53.15	1.24
4int	vinyl sulfone	Thr1	R	-145.02	1.36	S	-145.02	1.36
3ubb	phosphono-fluoridate	Ser201	R	-83.33	2.46	R	-78.00	1.45
3amh	β -lactam	Ser201		-112.85	1.27		-106.04	1.09
8est	isocoumarin	Ser195		-49.08	1.56		-48.81	0.38
linc	benzoxazinone	Ser195		-80.19	1.24		-77.25	1.02
1max	phosphonate	Ser195		-118.37	1.00		-118.37	1.00
1ldt	sulfonyl halide	Ser195		-37.88	0.49		-37.88	0.48
				mean:	1.74			1.36
				<2 Å:	71.4%			88.6%
				<1 Å:	20.0%			28.6%
				R/S ^c :	68.2%			

^aConfiguration of the electrophilic attachment point in the bound state for all prochiral ligands. ^bAll heavy atom RMSD values are calculated in angstroms from the unmodified X-ray structures. ^cPercentage of correctly predicted configurations of electrophilic attachment points in bound states.

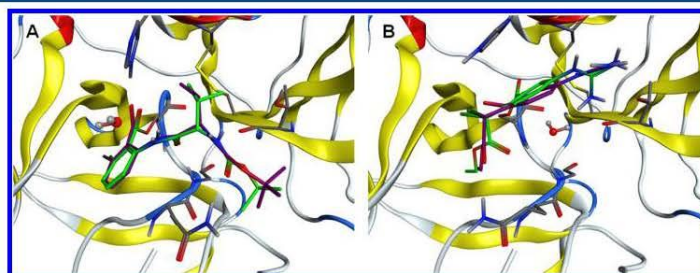


Figure 4. Successful binding mode prediction for ring opening reactions: A predicted binding mode of PDB-ID linc (benzoxazinone warhead, RMSD = 1.24 Å); B predicted binding mode of PDB-ID 8est (isocoumarin warhead, RMSD = 1.56 Å). Green: ligand carbon atoms as observed in the native X-ray structure. Purple: ligand carbon atoms of the predicted pose. Gray: receptor carbon atoms.

The RMSDs of the top scoring poses vary from 0.49 to 4.46 Å with a mean deviation of 1.74 Å (see Table 2). Using an RMSD cutoff of 2 Å for correctly predicted poses, DOCKTITE was able to predict experimentally determined native binding modes for 71.4% of the test set. The correct configuration (*R* or *S*) of the electrophilic attachment point for prochiral warheads was predicted in 68.2% of all cases. To check whether pose generation or the scoring function is the weaker link in the chain, we expanded our criterion of success to the top 10 predicted poses which increased the prediction rate of native binding modes to 88.6% with a mean RMSD of 1.36 Å.

Two examples for inherently difficult covalent docking tasks are shown in Figure 4. The difficulties with benzoxazinone and isocoumarin warheads arise because of the ring opening reaction upon ligand binding as shown in Figure 2 and are even more obvious with warheads containing leaving groups (see the Supporting Information). However, in both cases DOCKTITE predicted near-native binding modes with RMSD values of 1.24 Å (benzoxazinone, Figure 4A) and 1.56 Å (isocoumarin, Figure 4B), respectively.

Virtual Screening. The ability to screen large databases for putative covalent inhibitors in a manageable time scale on a common desktop computer was one of the key issues during the development of DOCKTITE. As prediction accuracy and speed often relate inversely, the user has to choose a compromise optimized for a specific project. To give an idea of how such a compromise could look like, the virtual screening validation was conducted twice using an accurate but slow force field refinement step as well as a very fast grid minimization refinement. The data set used consists of 7 active and 42 inactive 20S proteasome inhibitors from 7 warhead classes (aldehydes, ketones, α -keto amides, epoxyketones, β -lactams, boronates, and vinyl sulfones; see the Supporting Information). All ligands were docked into the structure of the bortezomib/20S proteasome complex (PDB-ID 2fl6, $\text{RMSD}_{\text{redocking}} = 1.24$ Å). That DOCKTITE is able to predict suitable binding modes for 20S proteasome inhibitors was additionally demonstrated in our pose prediction study using PDB-IDs: 4lqi and 4int (see Table 2).

The identification of active ligands within the library was realized using ligand efficiency ($\text{LE} = \text{London } dG \text{ score}/\text{number of heavy atoms}$) values derived from the consensus scoring approach, mentioned previously.³⁵ As a metric for screening quality, the area under curve (AUC) for the receiver operating characteristics (ROC) was used. In a perfect virtual screening, all active ligands would be scored on top of all inactive ligands with an AUC of 1 (see Figure 5, ideal curve). With a random selection of ligands the ROC shows a diagonal line and an AUC of 0.5 (see Figure 5, random curve).

Using the exhaustive mode (force field refinement, Figure 5, blue curve), the AUC of the virtual screening validation was 0.81, which shows that DOCKTITE has a very high predictive power of active ligands in the validation set. With this method, a database of 1000 ligands would approximately need 500–1000 h to complete using a common desktop workstation, which is in the time scale of most covalent docking programs, e.g., CovDock (1–3 h per ligand).⁹ In fast mode (grid refinement, Figure 5, red curve), the drop in predictivity is just marginal with an AUC of 0.79 but the gain in speed is enormous: a database of 1000 ligands now approximately needs only 3–5 h to complete. Furthermore, the early enrichment using the faster grid minimization refinement turned out to be superior to using the time-consuming forcefield refinement.

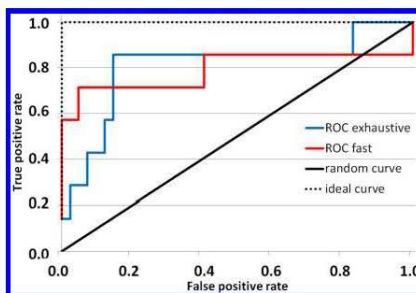


Figure 5. ROC for an exhaustive virtual screening, using force field refinement (blue curve, AUC = 0.81) and fast virtual screening, using grid minimization in the refinement step (red curve, AUC = 0.79).

It has to be noted, that a docking-based virtual screening of potential covalent inhibitors is an extraordinarily difficult task, because input databases generally consist of ligands bearing a variety of electrophilic warheads that differ in reactivity toward a nucleophilic receptor residue. Most scoring functions, also the ones used in DOCKTITE, do not score these reactivity patterns explicitly. As the reactivity of a specific electrophilic warhead is influenced by the rest of the molecule, such a scoring would require a time-consuming quantum mechanical calculation for each ligand, which is counterproductive to a fast virtual screening. Another problem is that IC_{50} or K_i values for irreversible or pseudoirreversible covalent inhibitors are time-dependent and therefore highly controlled by the experimental details.

Ranking of Congeneric Ligands. The validation of scoring power, or affinity prediction, of DOCKTITE was done with a congeneric series of 10 covalent c-Src kinase inhibitors, binding to Cys345 via a Michael acceptor mechanism.³⁴ The success critically depends on the scoring function used and is extremely error-prone due to simplifications of the underlying biophysics.¹⁴ Adequate calculation of entropic and desolvation effects or protein flexibility would benefit from computationally expensive methods like energy perturbation or MM-PBSA calculations. Furthermore, many covalent inhibitors show an irreversible or pseudoirreversible binding mode, thus, the experimental determination of IC_{50} values takes place in a nonequilibrium state and the apparent results are time-dependent.

However, it is expected that in one congeneric series of ligands with IC_{50} values determined under identical experimental conditions, the obtained ligand efficiency scores ($\text{LE} = \text{Affinity } dG \text{ score}/\text{number of heavy atoms}$) show a modest correlation with experimentally determined apparent binding affinities (pIC_{50}) and that DOCKTITE is of valuable use in the lead optimization process during drug development. In fact, the resulting calculated binding affinities displayed a significant correlation with the experimentally determined values resulting in a Pearson correlation coefficient of $\rho = 0.806$ ($R^2 = 0.649$, $p < 0.005$, Figure 6A). All ligands were docked into the X-ray structure of an acrylamide/c-Src kinase complex (PDB-ID: 2qlq, $\text{RMSD}_{\text{redocking}} = 2.73$ Å). Additional complexes for c-Src kinase inhibitors used in our pose prediction validation are PDB-IDs: 2hwo, 2hwp, and 3lok (see Table 2).

To further enhance the scoring power of DOCKTITE, the implementation of more precise scoring algorithms will be the

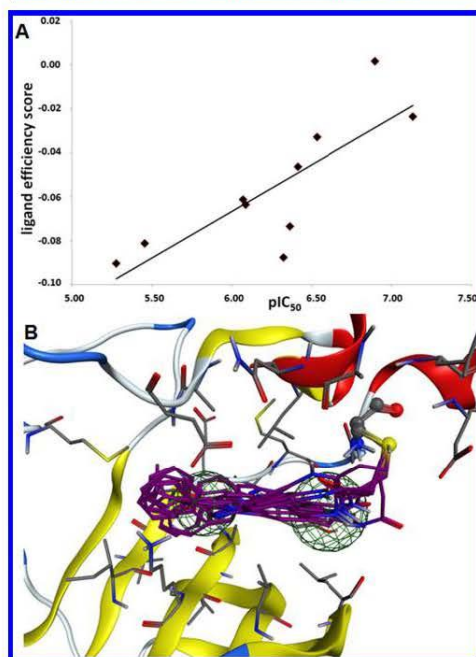


Figure 6. Ranking of congeneric ligands: A correlation plot of 10 c-Src kinase inhibitors docked in the X-ray structure of PDB-ID 2qlq ($\rho = 0.806$, $R^2 = 0.649$, $p < 0.005$); B obtained docking poses mapped on the utilized ph4 model. Purple: ligand carbon atoms. Gray: receptor carbon atoms. Ph4 features are represented as colored spheres: sulfur (yellow), oxygen (red), nitrogen (blue), carbon (gray), any heavy atom (green).

main focus in the future advancement and is under current investigation.

CONCLUSIONS

With DOCKTITE, we have developed a reliable, highly versatile, and user-friendly workflow for pose prediction and virtual screening of covalently bound ligands. Furthermore, it is possible to predict experimental binding affinities with good precision in convenient congeneric series of ligands. The warhead screening step is connected to the *MedChem Transformations* application and allows the user to implement any binding reaction in a very intuitive way. Twenty-one of the most common electrophilic warheads are already preinstalled. Because of the modular and open source nature of DOCKTITE, users may modify any step on their demands and may even outsource specific tasks to other software. The automated pharmacophore model is suitable to guide the docking simulation of any kind of ligand or receptor and can be refined as the knowledge of essential interactions increases during the drug development process. The novel consensus scoring approach used in DOCKTITE combines the strengths of the knowledge-based scoring function DSX and the empirical scoring functions implemented in MOE to give authentic information on near-native docking poses and their related

binding affinities. In addition to our pose prediction validation this combined strength becomes apparent in the additional validations of virtual screening performance, where DOCKTITE was able to differentiate between binders and nonbinders very precisely and the validation of scoring power, where the linear correlation of predicted with experimental binding affinities was significant. These characteristics enable DOCKTITE to be a valuable aide at all stages of a modern drug development process and clearly extend the already excellent software environment of MOE.

Software Download. All DOCKTITE SVL scripts including the source codes and graphical user interfaces are available free of charge from the *SVL Exchange* Web site <http://svlchemcomp.com>.

ASSOCIATED CONTENT

Supporting Information

Pose prediction results of 76 complexes used in the comparison with other docking programs (Table S1), complete list of implemented electrophilic warheads (Figure S1), structures of proteasome inhibitors used in virtual screening (Chart S1), and structures of c-Src kinase inhibitors used in ranking of congeneric ligands (Chart S2). This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: chrizz@chrizzweb.de. Phone: +49-6151 16 2047 (C.S.).

*E-mail: schmidt_boris@t-online.de. Phone: +49-6151 16 3075 (B.S.).

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), Germany (GrK1657, projects 2E and 3C).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Guido Kirsten, Dr. Kristina Grabowski, Dr. Barbara Sander (Chemical Computing Group), Dr. Paul Czodrowski (Merck KGaA), Dennis Bensinger, and Dr. Constantin Voss (TU Darmstadt) for their help and valuable comments.

ABBREVIATIONS

AUC, area under the curve; CPU, central processing unit; DSX, drug score extended; FDA, Federal Food and Drug Administration; ID, identification number; LE, ligand efficiency; MCT, MedChem Transformations; MM-PBSA, Molecular Mechanics Poisson–Boltzmann Surface Area; MOE, Molecular Operating Environment; n.a., not available; ρ , Pearson correlation coefficient; ph4, pharmacophore; RMSD, root of the mean square distance; pIC_{50} , $-\log(IC_{50})$; ROC, Receiver Operating Characteristics; SVL, Scientific Vector Language

REFERENCES

- (1) Vilar, S.; Cozza, G.; Moro, S. Medicinal chemistry and the Molecular Operating Environment (MOE): application of QSAR and molecular docking to drug discovery. *Curr. Top. Med. Chem.* **2008**, *8*, 1555–1572.

- (2) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; Lalonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, L. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.* **2005**, *49*, 5912–5931.
- (3) Kisselev, A. F.; van der Linden, W. A.; Overkleeft, H. S. Proteasome Inhibitors: An Expanding Army Attacking a Unique Target. *Chem. Biol.* **2012**, *19*, 99–115.
- (4) Mah, R.; Thomas, J. R.; Shafer, C. M. Drug discovery considerations in the development of covalent inhibitors. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 33–39.
- (5) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (6) Schröder, J.; Klinger, A.; Oellien, F.; Marhöfer, R. J.; Duszko, M.; Selzer, P. M. Docking-based virtual screening of covalently binding ligands: an orthogonal lead discovery approach. *J. Med. Chem.* **2013**, *56*, 1478–1490.
- (7) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- (8) Ouyang, X.; Zhou, S.; Su, C. T.; Ge, Z.; Li, R.; Kwok, C. K. CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. *J. Comput. Chem.* **2013**, *34*, 326–336.
- (9) Zhu, K.; Borrelli, K. W.; Greenwood, J. R.; Day, T.; Abel, R.; Farid, R. S.; Harder, E. Docking Covalent Inhibitors: A Parameter Free Approach To Pose Prediction and Scoring. *J. Chem. Inf. Model.* **2014**, *54*, 1932–1940.
- (10) London, N.; Miller, R. M.; Krishnan, S.; Uchida, K.; Irwin, J. J.; Eidam, O.; Gibold, L.; Cimermančič, P.; Bonnet, R.; Shoichet, B. K.; Taunton, J. Covalent docking of large libraries for the discovery of chemical probes. *Nat. Chem. Biol.* **2014**, *10*, 1066–1072.
- (11) DOCKovalent—Job Preparation Form. <http://covalent.docking.org/> (accessed December 13, 2014).
- (12) Mysinger, M. M.; Shoichet, B. K. Rapid Context-Dependent Ligand Desolvation in Molecular Docking. *J. Chem. Inf. Model.* **2010**, *50*, 1561–1573.
- (13) Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc.: Montreal, QC, Canada, 2013.
- (14) Neudert, G.; Klebe, G. DSX: A Knowledge-Based Scoring Function for the Assessment of Protein–Ligand Complexes. *J. Chem. Inf. Model.* **2011**, *51*, 2731–2745.
- (15) Santavy, M.; Labute, P. SVL: The Scientific Vector Language. <http://www.chemcomp.com/journal/svl.htm> (accessed October 7, 2014).
- (16) Daylight Theory Manual. <http://www.daylight.com/dayhtml/doc/theory/theory.smirks.html> (accessed October 20, 2014).
- (17) CTF File Formats. *Rxnfiles*; Symyx Technologies, Inc., 2007; Chapter 7.
- (18) Powers, J. C.; Asgian, J. L.; Ekici, Ö. D.; James, K. E. Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev.* **2002**, *102*, 4639–4750.
- (19) Gerber, P.; Müller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided Mol. Des.* **1995**, *9*, 251–268.
- (20) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198–210.
- (21) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (22) Halgren, T. A. Merck molecular force field. *J. Comput. Chem.* **1996**, *17*, 490–641.
- (23) Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins: Struct., Funct., Bioinf.* **2009**, *75*, 187–205.
- (24) Labute, P. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. *J. Comput. Chem.* **2008**, *29*, 1693–1698.
- (25) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.
- (26) Chen, X.; Lin, Y.; Liu, M.; Gilson, M. K. The Binding Database: data management and interface design. *Bioinformatics* **2002**, *18*, 130–139.
- (27) The Binding Database. <http://www.bindingdb.org> (accessed August 11, 2014).
- (28) Micale, N.; Ettari, R.; Lavecchia, A.; Di Giovanni, C.; Scarbaci, K.; Troiano, V.; Grasso, S.; Novellino, E.; Schirmeister, T.; Zappala, M. Development of peptidomimetic boronates as proteasome inhibitors. *Eur. J. Med. Chem.* **2013**, *64*, 23–34.
- (29) Geurink, P. P.; van der Linden, W. A.; Mirabella, A. C.; Gallastegui, N.; de Bruin, G.; Blom, A. E. M.; Voges, M. J.; Mock, E. D.; Florea, B. I.; van der Marel, G. A.; Driessen, C.; van der Stelt, M.; Groll, M.; Overkleeft, H. S.; Kisselev, A. F. Incorporation of non-natural amino acids improves cell permeability and potency of specific inhibitors of proteasome trypsin-like sites. *J. Med. Chem.* **2013**, *56*, 1262–1275.
- (30) Desvergne, A.; Genin, E.; Marechal, X.; Gallastegui, N.; Dufau, L.; Richy, N.; Groll, M.; Vidal, J.; Reboud-Ravaux, M. Dimerized Linear Mimics of a Natural Cyclopeptide (TMC-95A) Are Potent Noncovalent Inhibitors of the Eukaryotic 20S Proteasome. *J. Med. Chem.* **2013**, *56*, 3367–3378.
- (31) Kale, A. J.; Moore, B. S. Molecular Mechanisms of Acquired Proteasome Inhibitor Resistance. *J. Med. Chem.* **2012**, *55*, 10317–10327.
- (32) Geurink, P. P.; Liu, N.; Spaans, M. P.; Downey, S. L.; van den Nieuwendijk, A. M. C. H.; van der Marel, G. A.; Kisselev, A. F.; Florea, B. I.; Overkleeft, H. S. Incorporation of Fluorinated Phenylalanine Generates Highly Specific Inhibitor of Proteasome's Chymotrypsin-like Sites. *J. Med. Chem.* **2010**, *53*, 2319–2323.
- (33) Ryzdzewski, R. M.; Burrill, L.; Mendonca, R.; Palmer, J. T.; Rice, M.; Tahiramani, R.; Bass, K. E.; Leung, L.; Gjerstad, E.; Jang, J. W.; Pan, L. Optimization of subsite binding to the $\beta 5$ subunit of the human 20S proteasome using vinyl sulfones and 2-keto-1,3,4-oxadiazoles: Syntheses and cellular properties of potent, selective proteasome inhibitors. *J. Med. Chem.* **2006**, *49*, 2953–2968.
- (34) Blair, J. A.; Rauh, D.; Kung, C.; Yun, C.-H.; Fan, Q.-W.; Rode, H.; Zhang, C.; Eck, M. J.; Weiss, W. A.; Shokat, K. M. Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat. Chem. Biol.* **2007**, *3*, 229–238.
- (35) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430–431.

Supporting Information for

DOCKTITE - A Highly Versatile Step-by-Step Workflow for Covalent Docking and Virtual Screening in MOE.

Christoph Scholz^{†,}, Sabine Knorr^{||}, Kay Hamacher^{||} and Boris Schmidt^{†,*}*

[†]Clemens Schöpf - Institute of Organic Chemistry and Biochemistry, Technische Universität
Darmstadt, Alarich-Weiss-Str. 4, 64287 Darmstadt, Germany

^{||}Computational Biology & Simulation, Technische Universität Darmstadt,
Schnittspahnstraße 10, 64287 Darmstadt, Germany

Table of Content:

S2	I.	Pose prediction results of 76 covalent complexes.
S5	II.	Complete list of implemented electrophilic warheads.
S6	III.	Structures of proteasome inhibitors used in virtual screening.
S11	IV.	Structures used in ranking a congeneric series of c-Src kinase inhibitors.
S12	V.	References.

Table S1. Pose prediction results of 76 covalent complexes. All RMSD values in Å.

input complex				predicted pose			best pose in top 10	
PDB	Warhead	residue	R/S	score ^{DSX}	RMSD	R/S	score ^{DSX}	RMSD
2awz	Acrylamide	Cys366	R	-65.88	2.45	R	-62.29	1.13
2ax0	Acrylamide	Cys366	R	-85.06	1.00	R	-85.06	1.00
2ax1	Acrylamide	Cys366	R	-65.47	1.68	R	-57.92	1.20
2e14	Vinyl ketone	Cys166	S	-98.14	1.76	R	-98.14	1.76
2hwo	Acrylamide	Cys345	-	-87.63	1.94	-	-77.17	1.37
2hwp	Acrylamide	Cys345	-	-61.69	2.38	-	-61.69	2.38
2jiv	Acrylamide	Cys797	S	-187.30	1.74	R	-183.41	1.59
2qlq	Acrylamide	Cys345	R	-99.80	2.73	R	-89.07	1.13
3c9w	Vinylketone	Cys164	R	-102.06	1.08	R	-99.38	1.07
3ika	Acrylamide	Cys797	-	-23.00	4.46	-	-23.00	4.46
3lok	Acrylamide	Cys345	-	-75.21	1.60	-	-69.29	0.78
3oyp	Acrylamide	Cys159	-	-99.05	2.38	-	-88.98	2.35
3t9t	Acrylamide	Cys442	S	-132.14	1.39	R	-128.62	0.65
1b12	β-Lactam	Ser90	-	-82.17	2.29	-	-43.37	1.84
1bt5	β-Lactam	Ser70	-	-41.92	1.67	-	-40.78	1.07
1cef	β-Lactam	Ser62	-	-65.89	3.68	-	-55.50	3.10
1ceg	β-Lactam	Ser62	-	-59.89	2.80	-	-46.10	1.33
1fcm	β-Lactam	Ser61	-	-72.33	2.36	-	-71.95	2.36
1fcn	β-Lactam	Ser61	-	-76.52	4.53	-	-67.45	1.29
1fco	β-Lactam	Ser61	-	-86.84	2.55	-	-86.84	2.55
1fr6	β-Lactam	Ser64	-	-59.48	3.80	-	-53.75	3.59
1ghm	β-Lactam	Ser70	-	-83.71	1.15	-	-80.59	0.75
1ghp	β-Lactam	Ser70	-	-53.11	4.87	-	-46.74	0.91

1i5q	β -Lactam	Ser64	-	-66.59	2.25	-	-60.96	1.95
1iyp	β -Lactam	Ser70	-	-83.44	1.28	-	-83.44	1.28
1iyq	β -Lactam	Ser70	-	-74.85	2.74	-	-74.85	2.74
1kvm	β -Lactam	Ser64	-	-74.95	1.55	-	-74.79	1.43
1ll5	β -Lactam	Ser64	-	-43.83	2.72	-	-37.54	2.04
1ll9	β -Lactam	Ser64	-	-64.09	1.95	-	-62.01	1.94
1llb	β -Lactam	Ser64	-	-84.91	1.47	-	-82.89	1.44
1pw8	β -Lactam	Ser62	-	-72.85	3.58	-	-65.20	2.97
1pwd	β -Lactam	Ser62	-	-69.94	1.40	-	-65.16	0.87
1pwg	β -Lactam	Ser62	-	-88.80	1.82	-	-88.80	1.82
1qmf	β -Lactam	Ser337	-	-122.44	1.03	-	-113.75	0.88
1w8y	β -Lactam	Ser49	-	-70.79	1.25	-	-61.82	1.22
1ymx	β -Lactam	Ser70	-	-76.70	1.64	-	-76.70	1.64
2c5w	β -Lactam	Ser370	-	-101.88	1.26	-	-101.88	1.26
2ex6	β -Lactam	Ser62	-	-45.52	1.96	-	-37.53	0.87
2ex8	β -Lactam	Ser62	-	-51.21	1.39	-	-51.21	1.39
2ex9	β -Lactam	Ser62	-	-67.76	1.79	-	-60.52	0.59
2exa	β -Lactam	Ser62	-	-33.41	3.60	-	0.76	1.86
2exb	β -Lactam	Ser62	-	-95.11	4.68	-	-73.53	3.75
2j8y	β -Lactam	Ser61	-	-61.76	2.26	-	-56.47	1.25
2jbf	β -Lactam	Ser61	-	-41.83	4.62	-	-36.86	4.03
2vgj	β -Lactam	Ser49	-	-75.53	1.38	-	-75.53	1.38
2wkh	β -Lactam	Ser67	-	-55.88	2.00	-	-54.88	1.05
2xd1	β -Lactam	Ser460	-	-105.24	2.65	-	-103.41	2.04
2z2m	β -Lactam	Ser337	-	-144.76	1.91	-	-144.76	1.91
2zc3	β -Lactam	Ser337	-	-64.43	0.58	-	-64.43	0.58
2zc4	β -Lactam	Ser337	-	-73.44	1.54	-	-67.39	1.44

2zc5	β -Lactam	Ser370	-	-86.69	3.21	-	-62.93	2.05
2zc6	β -Lactam	Ser370	-	-83.05	3.78	-	-65.35	1.42
2zd8	β -Lactam	Ser70	-	-31.26	1.11	-	-28.84	1.00
2zq9	β -Lactam	Ser70	-	-87.54	0.48	-	-87.54	0.48
2zqa	β -Lactam	Ser70	-	-79.27	1.60	-	-59.88	1.45
2zqc	β -Lactam	Ser70	-	-71.82	3.62	-	-71.82	3.62
2zqd	β -Lactam	Ser70	-	-41.29	2.89	-	-41.29	2.89
3a3e	β -Lactam	Ser69	-	-81.10	2.65	-	-72.30	1.80
3a3f	β -Lactam	Ser69	-	-97.76	2.96	-	-97.76	2.96
3a3i	β -Lactam	Ser69	-	-57.86	5.08	-	-54.37	2.49
3beb	β -Lactam	Ser44	-	-67.70	5.14	-	-44.76	1.83
3bec	β -Lactam	Ser44	-	-17.02	0.90	-	-17.02	0.90
3dwz	β -Lactam	Ser84	-	-64.44	4.08	-	-58.39	1.95
3iqa	β -Lactam	Ser84	-	-65.12	2.77	-	-65.12	2.77
3ita	β -Lactam	Ser40	-	-26.09	1.88	-	-14.03	1.39
3kgo	β -Lactam	Ser70	-	-73.92	1.39	-	-65.73	1.26
3ly4	β -Lactam	Ser70	-	-72.88	1.86	-	-59.26	1.66
3m6b	β -Lactam	Ser84	-	-78.85	1.43	-	-78.85	1.43
3m6h	β -Lactam	Ser84	-	-89.19	2.07	-	-76.56	1.45
3mzd	β -Lactam	Ser44	-	-109.30	2.40	-	-102.46	1.60
3mze	β -Lactam	Ser44	-	-81.33	2.39	-	-72.05	1.97
3mzf	β -Lactam	Ser44	-	-48.31	4.71	-	-42.64	0.78
3n8s	β -Lactam	Ser84	-	-57.68	2.91	-	-57.52	2.91
3ocl	β -Lactam	Ser294	-	-101.16	1.92	-	-93.84	1.28
3ocn	β -Lactam	Ser294	-	-93.06	2.78	-	-70.43	1.67
3pbo	β -Lactam	Ser294	-	-121.61	3.53	-	-89.29	1.99
Mean:					2.40			1.74

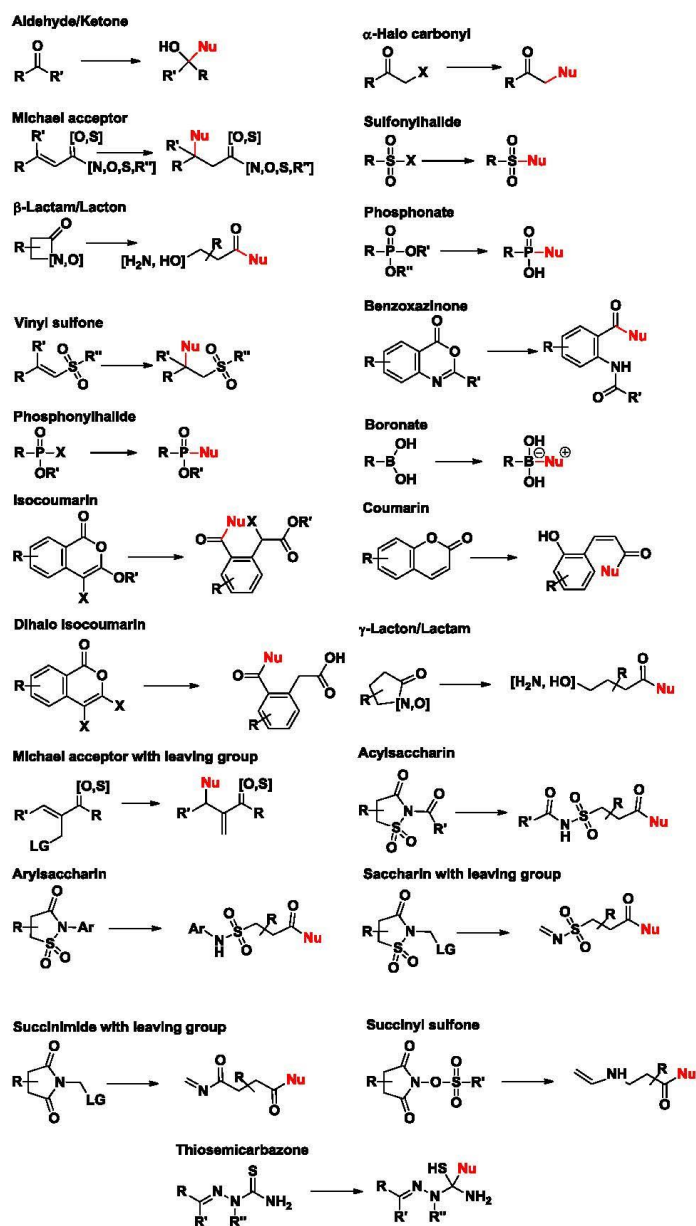
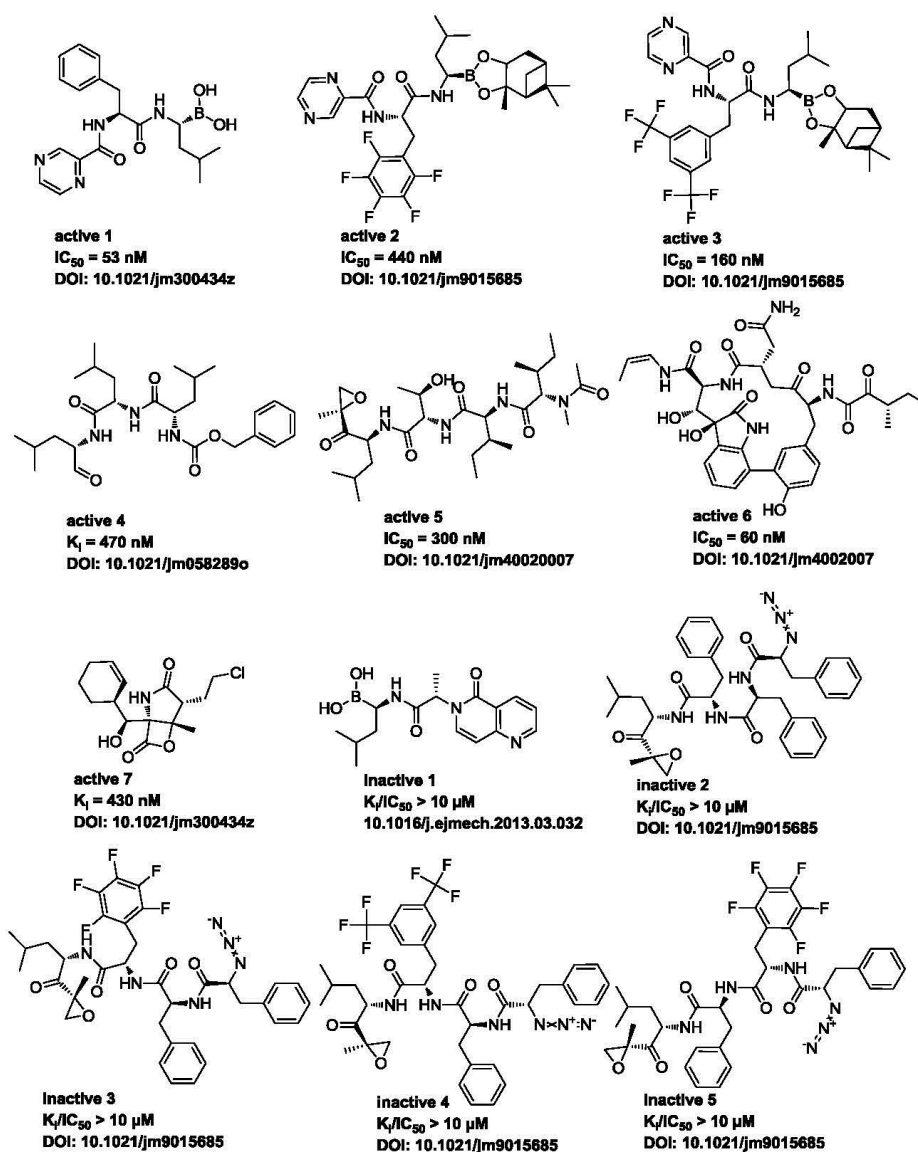
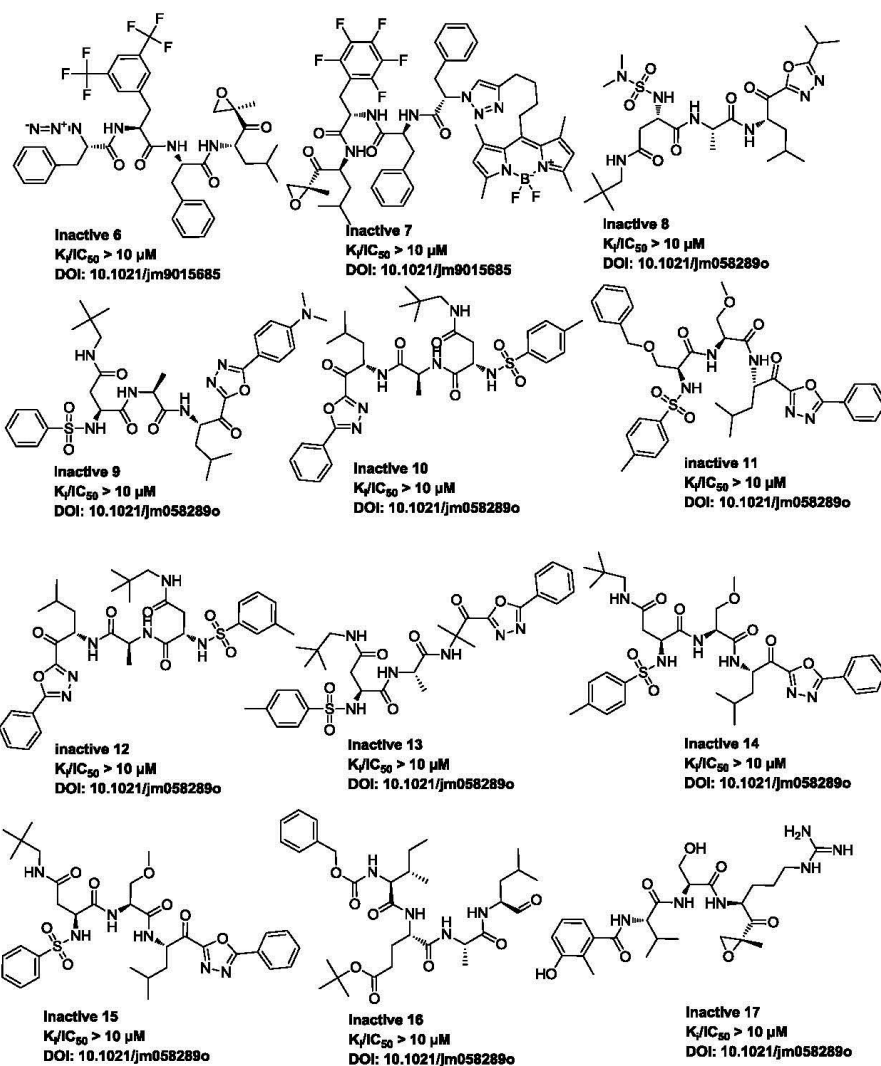
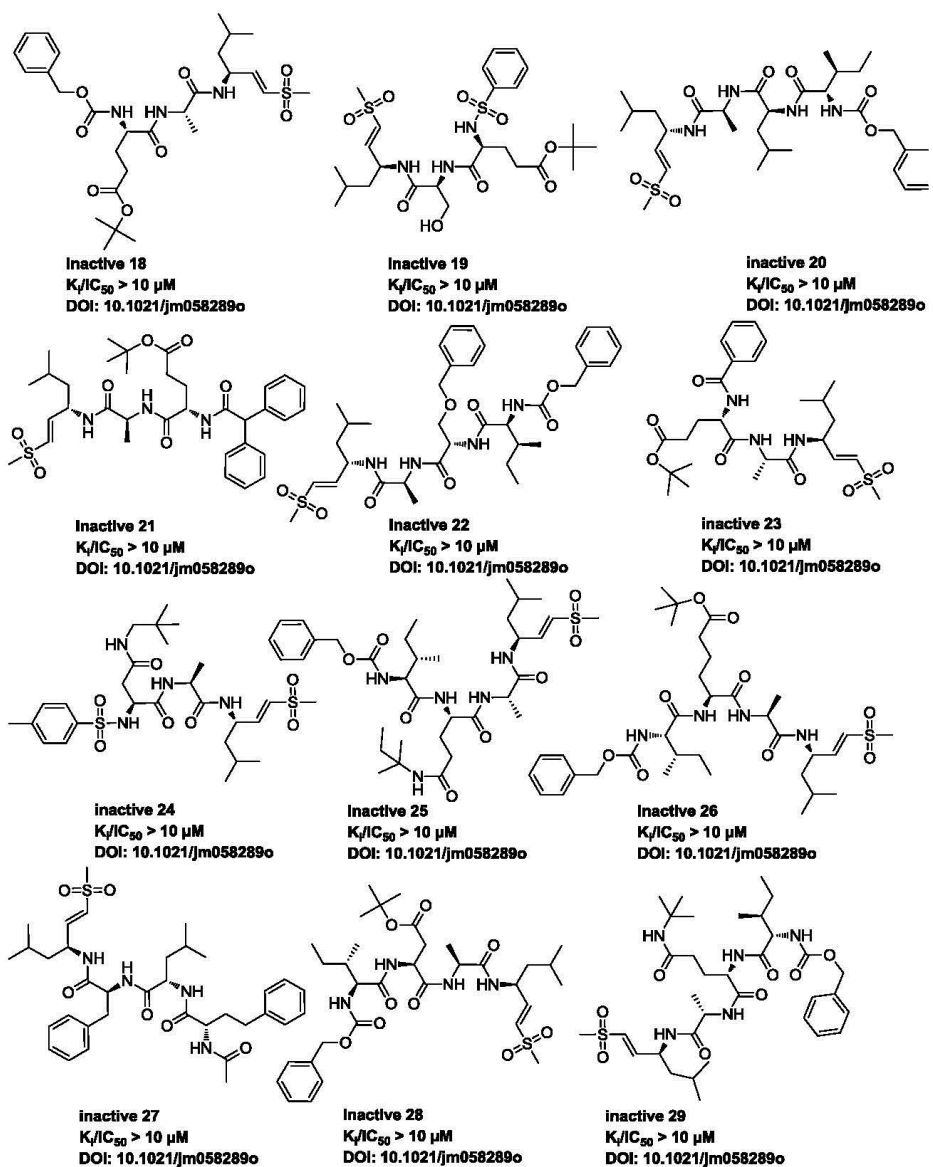


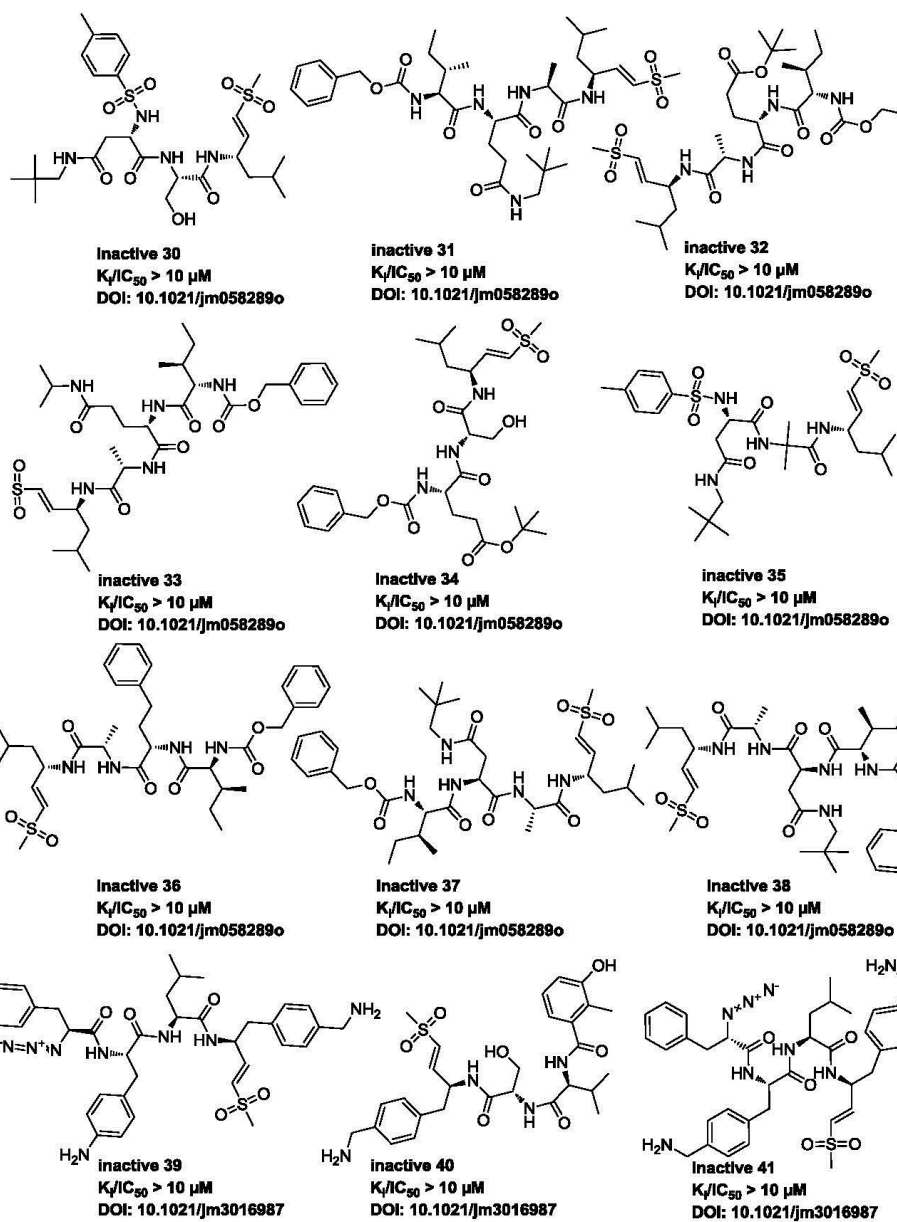
Figure S1. Complete list of implemented electrophilic warheads in DOCKTITE.

Chart S1. Compounds used in virtual screening.¹⁻⁶









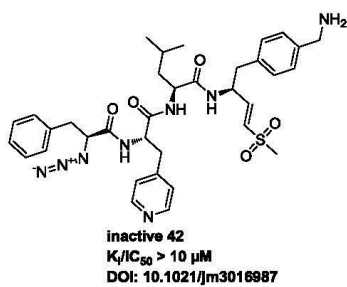
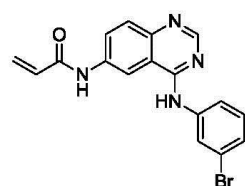
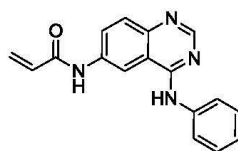


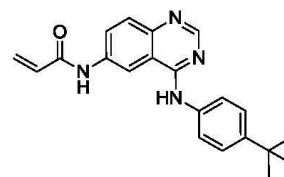
Chart S2: Compounds used in ranking a congeneric series of c-Src kinase inhibitors.⁷



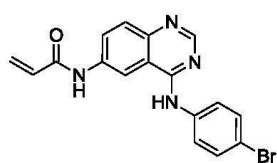
acrylamide 1
 $IC_{50} = 73 \text{ nM}$



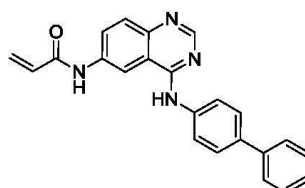
acrylamide 2
 $IC_{50} = 126 \text{ nM}$



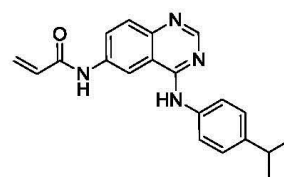
acrylamide 3
 $IC_{50} = 5.29 \mu\text{M}$



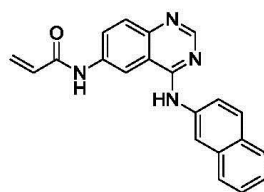
acrylamide 4
 $IC_{50} = 294 \text{ nM}$



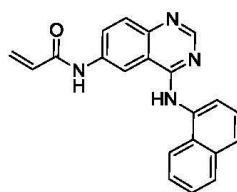
acrylamide 5
 $IC_{50} = 387 \text{ nM}$



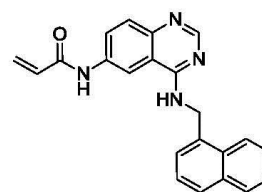
acrylamide 6
 $IC_{50} = 3.50 \mu\text{M}$



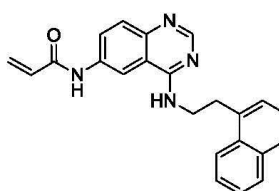
acrylamide 7
 $IC_{50} = 852 \text{ nM}$



acrylamide 8
 $IC_{50} = 820 \text{ nM}$



acrylamide 9
 $IC_{50} = 477 \text{ nM}$



acrylamide 10
 $IC_{50} = 436 \text{ nM}$

REFERENCES

- (1) Rydzewski, R. M.; Burrill, L.; Mendonca, R.; Palmer, J. T.; Rice, M.; Tahilramani, R.; Bass, K. E.; Leung, L.; Gjerstad, E.; Janc, J. W.; Pan, L., Optimization of subsite binding to the $\beta 5$ subunit of the human 20S proteasome using vinyl sulfones and 2-keto-1,3,4-oxadiazoles: Syntheses and cellular properties of potent, selective proteasome inhibitors. *J. Med. Chem.* **2006**, *49*, 2953-2968.
- (2) Geurink, P. P.; Liu, N.; Spaans, M. P.; Downey, S. L.; van den Nieuwendijk, A. M. C. H.; van der Marel, G. A.; Kisselev, A. F.; Florea, B. I.; Overkleeft, H. S., Incorporation of Fluorinated Phenylalanine Generates Highly Specific Inhibitor of Proteasome's Chymotrypsin-like Sites. *J. Med. Chem.* **2010**, *53*, 2319-2323.
- (3) Kale, A. J.; Moore, B. S., Molecular Mechanisms of Acquired Proteasome Inhibitor Resistance. *J. Med. Chem.* **2012**, *55*, 10317-10327.
- (4) Desvergne, A.; Genin, E.; Marechal, X.; Gallastegui, N.; Dufau, L.; Richy, N.; Groll, M.; Vidal, J.; Reboud-Ravaux, M., Dimerized Linear Mimics of a Natural Cyclopeptide (TMC-95A) Are Potent Noncovalent Inhibitors of the Eukaryotic 20S Proteasome. *J. Med. Chem.* **2013**, *56*, 3367-3378.
- (5) Geurink, P. P.; van der Linden, W. A.; Mirabella, A. C.; Gallastegui, N.; de Bruin, G.; Blom, A. E. M.; Voges, M. J.; Mock, E. D.; Florea, B. I.; van der Marel, G. A.; Driessen, C.; van der Stelt, M.; Groll, M.; Overkleeft, H. S.; Kisselev, A. F., Incorporation of non-natural amino acids improves cell permeability and potency of specific inhibitors of proteasome trypsin-like sites. *J. Med. Chem.* **2013**, *56*, 1262-1275.
- (6) Micale, N.; Ettari, R.; Lavecchia, A.; Di Giovanni, C.; Scarbaci, K.; Troiano, V.; Grasso, S.; Novellino, E.; Schirmeister, T.; Zappala, M., Development of peptidomimetic boronates as proteasome inhibitors. *Eur. J. Med. Chem.* **2013**, *64*, 23-34.
- (7) Blair, J. A.; Rauh, D.; Kung, C.; Yun, C.-H.; Fan, Q.-W.; Rode, H.; Zhang, C.; Eck, M. J.; Weiss, W. A.; Shokat, K. M., Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat. Chem. Biol.* **2007**, *3*, 229-238.

3.4 Entwicklung einer optimierten Syntheseroute zu Methoxy-XO4-basierten Fluoreszenzfarbstoffen zur Diagnose der Alzheimer-Demenz

Der Inhalt dieses Kapitels wurde bereits veröffentlicht:

Alexander Boländer, Daniel Kieser, Christoph Scholz, Roland Heyny-von Haußen, Gerhard Mall, Valérie Goetschy, Christian Czech, Boris Schmidt, "Synthesis of methoxy-XO4 derivatives and their evaluation in Alzheimer's disease pathology", *Neurodegenerative Diseases* **2014**, 13, 209-213.

Mit freundlicher Genehmigung des Verlags Karger Publishers. Nummerierungen wie im Manuskript.

Die niedermolekulare Fluoreszenzsonde Methoxy-XO4 **7a** ist in der Lage, die Blut/Hirn-Schranke zu überwinden und bindet dort an die AD-spezifischen Ablagerungen aus APs und NFTs (siehe Abschnitt 1.3.3). Einer Derivatisierungsstudie zur Optimierung metabolischer Parameter sowie der Selektivität gegenüber NFTs stand die etablierte dreistufige Syntheseroute mit aufwendigen Reinigungsschritten und relativ schlechten Ausbeuten im Wege.

Knackpunkt der Syntheseroute ist die radikalische Bromierung mit einer Ausbeute von nur 19% im ersten Schritt aufgrund auftretender Mehrfachbromierungen. Diese Problematik konnte durch die neu entwickelte vierstufige Syntheseroute umgangen und somit die Gesamtausbeute mit einer Erhöhung von 19,4% auf 37,4% nahezu verdoppelt werden. Mittels der neu entwickelten Syntheseroute konnten elf Derivate von Methoxy-XO4 **7a** hergestellt und histologisch sowie in einem Verdrängungs-Assay auf ihre Fähigkeit, selektiv an APs und NFTs zu binden, untersucht werden. Es konnten hochaffine Fluoreszenzsonden erhalten werden, von denen das 2,6-Dibrom-4-Hydroxyderivat **7g** eine exzellente Selektivität für NFTs und das Boronsäurederivat **7l** eine ausgeprägte Selektivität für APs zeigt. Aufgrund der besseren Korrelation erster Symptome der AD mit der NFT-Konzentration in Hirn und Retina ist besonders die Fluoreszenzsonde **7g** von herausragender Bedeutung für eine Weiterentwicklung der Fluoreszenzmikroskopie für die Diagnose der AD.

Beitrag C. Scholz: Synthese der Zwischenstufe **3** und Methoxy-XO4 **7a**.

Synthesis of Methoxy-X04 Derivatives and Their Evaluation in Alzheimer's Disease Pathology

Alexander Boländer^a Daniel Kieser^a Christoph Scholz^a
Roland Heyny-von Haußen^b Gerhard Mall^b Valérie Goetschy^c
Christian Czech^c Boris Schmidt^a

^aClemens Schöpf-Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, and

^bCenter of Pathology, Klinikum Darmstadt, Darmstadt, Germany; ^cF. Hoffmann-La Roche AG, Basel, Switzerland

Key Words

Alzheimer's disease · Fluorescence imaging · Tau · Amyloid- β · Histology

Abstract

Background: Alzheimer's disease is characterized by two notorious protein aggregates in the brain: extracellular senile plaques mainly consisting of amyloid- β peptides and tau-protein-derived intracellular paired helical filaments. The diagnosis of Alzheimer's disease is impaired by insufficient sensitivity and specificity of diagnostic methods to visualize these pathological hallmarks over all disease stages.

Objective: The established fluorescence marker methoxy-X04 stains plaques, tau tangles and amyloid-derived angiopathies with good specificity, yet it is limited by slow elimination in vivo. Since the need for new markers is high, we prepared methoxy-X04 derivatives and evaluated their potential as imaging agents in Alzheimer's disease pathology.

Methods and Results: In this study, we describe an improved synthesis for methoxy-X04 and its derivatives and their affinity determination for the respective protein targets by immunohistology and a displacement assay. **Conclusion:** This resulted in the identification of new derivatives of methoxy-X04 with improved binding affinity.

© 2013 S. Karger AG, Basel

Introduction

Alzheimer's disease (AD) is the most common form of dementia. With symptoms such as memory loss, disorientation, language and cognitive impairment, it is still an incurable and lethal disease accounting for the majority of cases of dementia diagnosed over the age of 60 [1]. The histopathology of AD is characterized by two notorious protein aggregates in the brain: extracellular senile plaques consisting of amyloid- β (A β) peptides and intracellular tau-protein-derived neurofibrillary tangles (NFTs) [2].

The overproduction and aggregation of A β in the brain represents one major cause of AD, which can be traced back to the misprocessing of the much larger A β precursor protein APP, prompted by two different enzymes, β - and γ -secretase. They release short A β peptides from its precursor, which form ordered oligomeric and fibrillar secondary structures rich in cross- β -sheets, clumping together and forming larger plaques. The amy-

A.B. and D.K. contributed equally to this work.

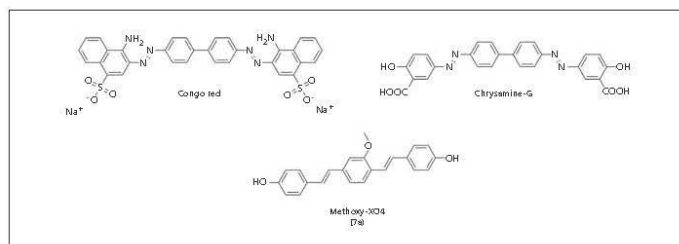
KARGER

E-Mail karger@karger.com
www.karger.com/ndd

© 2013 S. Karger AG, Basel
1660–2854/13/0134–0209\$38.00/0

Prof. Dr. Boris Schmidt
Clemens Schöpf-Institute of Organic Chemistry and Biochemistry
Technische Universität Darmstadt
Petersenstrasse 22, DE-64287 Darmstadt (Germany)
E-Mail schmidt_boris@t-online.de

Fig. 1. Fluorescent markers for the pathological hallmarks in AD: Congo red, Chrysamine-G and methoxy-X04 (7a).



loid hypothesis assigns this A β accumulation to be the primary event in the development of AD pathology [3]. A β -directed in vivo imaging and post mortem histology have shown that plaques start to accumulate 10–15 years before symptoms emerge and there are reports for amyloid plaque load at autopsy, yet no signs of cognitive impairment during lifetime [4]. These findings and other limitations restrict the diagnostic power of the first FDA-approved plaque-directed probe florbetapir to rule out AD, but not to diagnose it [5].

NFTs consist of paired helical filaments of abnormally hyperphosphorylated tau protein [6]. In healthy neurons, six highly soluble tau isoforms modulate the stabilization of axonal microtubules by their interaction with tubulin. The stabilizing ability of tau protein is lost after hyperphosphorylation by several serine/threonine and tyrosine protein kinases. Hyperphosphorylated tau has a high tendency to pair with other threads of tau-forming intracellular tangles, which lead to fatal neurodegeneration. In contrast to A β plaques, the correlation between tau fibrils and the timing of symptom onset and progression of AD seems to be far more accurate [4]. The visualization of both kinds of aggregates in the living patient may provide the desired diagnostic method to enhance in vivo diagnosis of AD. Currently there is no simple, inexpensive and definite method available to selectively detect A β plaques and NFTs in the brain of a living patient, which makes it difficult to diagnose AD distinctively and to differentiate it from other forms of dementia [7].

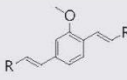


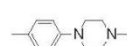







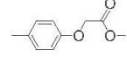
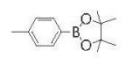
Tau-targeting probes have emerged more slowly than A β -directed probes. This is partly due to an overwhelming interest in A β , partly due to tau's essential role in maintaining cell viability and the lack of essential models of tau aggregation truly reflecting the pathology in humans. By contrast, multiple efforts have been made to develop A β -labeling markers for imaging techniques such as single photon emission tomography, positron emission tomography or other applications like fluores-

cence spectroscopy or microscopy [8]. One of the most prominent fluorescent markers is methoxy-X04 (7a; see table 1 for 7a and the synthesized derivatives 7b–l), a Congo red and Chrysamine-G-derived bisphenolic distyrylbenzene derivative (fig. 1) [9]. 7a was reported to stain amyloid plaques, NFTs and cerebrovascular amyloid by post mortem histochemical analysis of human brain tissue. Individual plaques in PS1/APP mice were monitored in vivo over months by multiphoton microscopy after intravenous injection of 7a [10]. However, this ligand seems inadequate to differentiate clearly between A β and tau deposits. This differentiation empowers the presymptomatic diagnostic potential of each biomarker and their correlation with disease progression and neurodegeneration. 7a displays a very slow clearance in mice and persists for weeks after systemic administration. This may be due to its rather high lipophilicity and restricts clinical application in humans [10]. Therefore, improved markers are needed which provide selective labeling of tau fibrils and amyloid plaques and display improved physicochemical properties. There are very few brain-permeating probes which exploit the tau load of the brain as a biomarker for the severity of dementia in AD [6]. Several of these properties, i.e. the potential to pass the blood brain barrier, the unspecific binding to white brain matter and the clearance after systemic administration, can be correlated to the lipophilicity of the probe (given by the logP value), its topological polar surface area (tPSA) and the molecular weight. Thus, we selected these cheminformatic properties as surrogate parameters for the early development.

Materials and Methods

Since both A β plaques and tau fibrils exhibit a secondary structure rich in β -pleated regions, yet differ in their primary sequence, this common structural motive suggests utilizing an established β -sheet binding moiety as starting point [2]. We selected 7a as the

Table 1. 7a and the synthesized derivatives 7b-l with determined affinities to the aggregated proteins (tau and A β) in a thiazin red displacement assay and spectroscopic and chemical properties data

7		IC ₅₀ ^a , nM		cLogP ^b	abs _{max} , nm	em _{max} , nm	tPSA ^c Å ²	MW
		aggregated tau	aggregated A β					
a		246	140	6.11	344	450	49.69	344.40
b		1.8	13	6.86	359	430	15.71	398.54
c		366	120	6.37	352	445	22.19	508.70
d		inactive	inactive	3.61	368	430	33.95	314.38
e		35	24	5.42	326	441	33.95	383.27
f		inactive	inactive	7.95	337	438	9.23	470.20
g		66	1,597	6.39	375	440	90.15	534.19
h		6,454	2,535	6.6	368	428	9.23	348.39
i		inactive	inactive	6.92	369	433	9.23	384.37
j		294	1,020	5.05	396	441	90.15	412.38
k		6.4	3.7	5.09	350	437	80.29	488.53
l		94	18	n.d.	372	445	46.15	564.33

abs_{max} = Absorption maximum; em_{max} = emission maximum; MW = molecular weight.
^a IC₅₀ values were determined in a thiazin red displacement assay as duplicates of replicate experiments (see online suppl. material for further details) [13].

^b Calculated cLogP values (cLogP values were determined using the OSIRIS property explorer, <http://www.organic-chemistry.org/prog/peo/>).

^c Calculated tPSA values (tPSA values were determined using ChemBioDraw Ultra 11.0).

Fig. 2. Reagents and conditions. **a** NBS, AIBN, CCl₄, reflux; 19%. **b** Triethyl phosphate; 97%. **c** KMnO₄, H₂O, 100°C, 12 h; 63%. **d** (1) SOCl₂, MeOH, 0°C, 2 h, (2) K₂CO₃, reflux, 3 h; 92%. **e** LiAlH₄, THF; 75%. **f** (1) SOCl₂, CH₂Cl₂, (2) triethyl phosphate; 86%. **g** KO^tBu, aromatic aldehyde, DMF; 68%. NBS = N-Bromsuccinimid; AIBN = alpha-isobutyronitrile; DMF = dimethylformamide; THF = tetrahydrofuran.

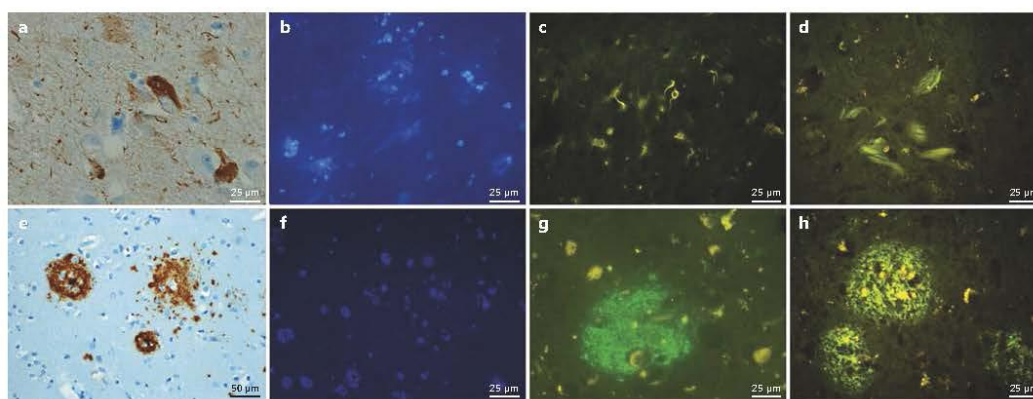
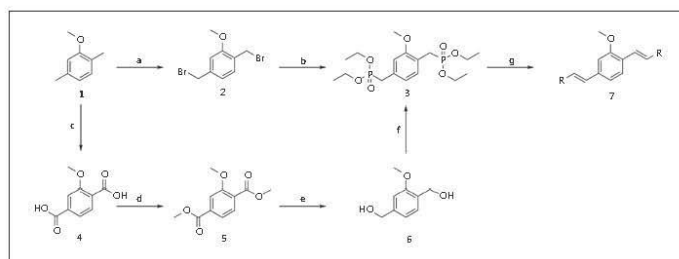


Fig. 3. Immunohistological staining of NFTs (a) and amyloid plaques (e) and histochemical staining of both kinds of deposits with 7a (b, f), 7b (c, g) and 7g (d, h).

lead compound because it allows derivatization of the core structure. Different moieties of variable size and electron-donating or electron-withdrawing properties were added as flanks to the central anisole core (fig. 2; numbers in the following paragraph refer to this figure).

The Wadsworth-Emmons phosphonate (3) was obtained by bromination of 2-methoxy-1,4-dimethylbenzene (1) and subsequent phosphorylation. The bromination of 1 by a radical mechanism results in overbromination, which reduces yield and requires the laborious purification of a lachrymator. Therefore, we took a different approach: again starting from 2-methoxy-1,4-dimethylbenzene (1) we oxidized it to the diacid (4), which was followed by esterification and subsequent reduction to yield the diol (6). Compound 6 was halogenated in situ and directly converted to the Wadsworth-Emmons phosphonate (3). The overall yield of 37.4% of this 'one additional step' synthesis compares to the 19.4% we obtained by the shorter route and avoids purification of noxious intermediates. Compound 7a was prepared by a Wadsworth-Em-

mons reaction of phosphonate (3) and 4-hydroxybenzaldehyde, favoring the formation of the trans-stilbene product in a multi-gram quantity. This step was achieved without methoxymethyl protection of the phenol group of 4-hydroxybenzaldehyde. By the use of various aromatic aldehydes, we prepared a series of derivatives (7), which are shown in table 1.

In order to evaluate the staining potential of these dyes, we applied them to human brain tissue affected by AD pathology (tissue specimens were obtained with patients' informed consent in the course of regular clinical autopsies performed at the department of pathology, Klinikum Darmstadt). The subsequent analysis by fluorescence microscopy demonstrated a good-to-excellent staining of either tau, Aβ or both kinds of deposits at high-contrast, low-background staining and high photostability, which was in accord with the immunohistological control staining. The antibodies anti-PHF tau clone AT8 mAb for tau deposits and amyloid A4 for Aβ plaques were used for the immunohistological staining (fig. 3).

Results and Discussion

The dimethylamine derivative (7b) showed the best staining results for both tau and A β . First hints at a potential ability to discriminate between both kinds of protein deposits were observable for 7g, showing better staining results for tangles. The derivative 7l labels plaques more intensely than NFTs. However, histological in vitro staining experiments provide only little information on absolute affinities and are not directly comparable with respect to prospective in vivo imaging experiments. Thus, we tested the compounds in a thiazin red displacement assay to determine their respective affinity to aggregated tau and A β_{40} (see online suppl. material for further details; www.karger.com/doi/10.1159/000351436). The required concentrations of the tested ligands to replace thiazin red from half of the present binding sites are tabulated as IC₅₀ values in table 1. The IC₅₀ values confirmed the trends of the qualitative in vitro staining experiments. Several compounds have an improved ability to discriminate between tau and A β in vitro and show generally higher affinities compared to the reference compound 7a. The minimum selectivity in bonding to the two proteins has to be elucidated in vivo. Observations of in vitro assays concluded that a 20- to 50-fold higher selectivity towards the desired target may be sufficient [11]. According to this hypothesis, 7g holds potential for a tau selective

marker in vivo. 7g has nearly the same lipophilicity as the reference compound 7a. The tPSA value (90.15 Å²) and molecular weight (534.90 g/mol) of 7g is high in comparison to 7a, yet still in a tolerable range for brain-permeating probes, which ideally possess a tPSA value of <80 Å² and a molecular weight of <500 g/mol [12].

Conclusion

We have developed a new synthesis for methoxy-X04 (7a), which enables rapid derivatization of the distyrylbenzene scaffold by the introduction of variable moieties as aromatic flanks. These derivatizations resulted in fluorescent molecules with improved binding properties to AD-associated NFTs and amyloid plaques, which was demonstrated in histochemical staining experiments on human brain tissue and by a thiazin red displacement assay in vitro. One marker shows a 24-fold higher affinity for NFTs compared to A β plaques and thus holds potential to be profiled as a tau-selective ligand in vivo.

Acknowledgments

This work was supported in part by the Bundesministerium für Bildung und Forschung, Germany (13N10636).

References

- Wimo A, Winblad B, Aguero-Torres H, von Strauss E: The magnitude of dementia occurrence in the world. *Alzheimer Dis Assoc Disord* 2003;17:63–67.
- Jakob-Roetne R, Jacobsen H: Alzheimer's disease: from pathology to therapeutic approaches. *Angew Chem Int Ed Engl* 2009;48:3030–3059.
- Hardy J, Selkoe DJ: The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 2002;297:353–356.
- Gravitz L: Drugs: a tangled web of targets. *Nature* 2011;475:S9–S11.
- Garber K: First FDA-approved beta-amyloid diagnostic hits the market. *Nat Biotech* 2012;30:575.
- Medeiros R, Baglietto-Vargas D, LaFerla FM: The role of tau in Alzheimer's disease and related disorders. *CNS Neurosci Ther* 2011;17:514–524.
- Herholz K, Ebmeier K: Clinical amyloid imaging in Alzheimer's disease. *Lancet Neurol* 2011;10:667–670.
- Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, Bergstrom M, Savitcheva I, Huang G-F, Estrada S, Ausen B, Debnath ML, Barletta J, Price JC, Sandell J, Lopresti BJ, Wall A, Koivisto P, Antoni G, Mathis CA, Langstrom B: Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol* 2004;55:306–319.
- Klunk WE, Bacskaï BJ, Mathis CA, Kajdasz ST, McLellan ME, Frosch MP, Debnath ML, Holt DP, Wang Y, Hyman BT: Imaging A β plaques in living transgenic mice with multiphoton microscopy and methoxy-X04, a systemically administered Congo red derivative. *J Neuropathol Exp Neurol* 2002;61:797–805.
- Burgold S, Bittner T, Dorostkar Mario M, Kieser D, Fuhrmann M, Mitteregger G, Kretschmar H, Schmidt B, Herms J: In vivo multiphoton imaging reveals gradual growth of newborn amyloid plaques over weeks. *Acta Neuropathol* 2011;121:327–335.
- Schafer Kelsey N, Kim S, Matzavinos A, Kuret J: Selectivity requirements for diagnostic imaging of neurofibrillary lesions in Alzheimer's disease: a simulation study. *Neuroimage* 2012;60:1724–1733.
- Schmidt B, Kieser D, Bolaender A, Herms J, Heyny-von Haussen RH: Fluorescent ligands for protein aggregates and their use in the diagnosis of neurodegenerative diseases at the retina; Klinikum Darmstadt GmbH, Germany; Ludwig-Maximilians-Universität München; Technische Universität Darmstadt 2012, p 118.
- Bolaender A, Kieser D, Voss C, Bauer S, Schoen C, Burgold S, Bittner T, Hoelzer J, Heyny-von Haussen R, Mall G, Goetschy V, Czech C, Knust H, Berger R, Herms J, Hilger I, Schmidt B: Bis(arylvinyl)pyrazines, pyrimidines, and pyridazines as imaging agents for tau fibrils and β -amyloid plaques in Alzheimer's disease models. *J Med Chem* 2012;55:9170–9180.

Supplementary Material for

**Synthesis of Methoxy-XO4 Derivatives and their Evaluation in
Alzheimer's Disease Models**

Alexander Boländer^{a†}, Daniel Kieser^{a†}, Christoph Scholz^a, Roland Henyny-von Haußen^b,
Gerhard Mall^b, Valérie Goetschy^c, Christian Czech^c and Boris Schmidt^{a*}

^a*Clemens Schöppf-Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, 64287 Darmstadt, Germany*

^b*Center of Pathology, Klinikum Darmstadt, 64287 Darmstadt, Germany*

^c*F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland*

[†]*both authors contributed equally to this work, both first authors.*

Tissue Probe Preparation

Human hippocampal tissue fixed in 10% buffered formalin solution and embedded in paraffin was cut to 4 μm sections in a microtome and mounted on standard microscope slides. Subsequent deparaffinization was carried out in xylol (15 min), ethanol 100 % (10 min), ethanol 96% (10 min) and ethanol 70% (10 min). Until subsequent processing the slides were stored in water.

Immunohistochemical Staining

Immunohistochemical staining was performed on 4 μm sections by using a Ventana Benchmark automated stainer (Ventana, Tuscon, AZ) applying the antibodies anti-PHF-Tau clone AT8 mAb (Thermo Scientific Pierce Protein Research Products, Rockford, IL, USA), TAU Ab-3 (Neomarkers, Fremont, CA) and Amyloid A4 (BAM10, Sigma, St. Louis, MO) and the Ultraview Universal DAB Detection Kit (Ventana, Tuscon, AZ).

Histochemical Staining with Fluorescent Dyes

The staining procedure comprises the following steps of brain section treatment: nucleus staining with acidic haemalm (10 min), blue annealing by water wash-up (5 min), application of the dye solution (2.5 mM) (10 min), ethanol wash-up, water-bath (5 min), 1% acetic acid-bath (15 min), water wash-up and covering with standard glass covers.

Light Microscopy and Fluorescence Microscopy

4 μm sections stained with fluorescence probes were analysed by indirect fluorescence microscopy using an Axioskop microscope with a HBO100 fluorescence illuminator (Carl Zeiss, Oberkochen, Germany) with the bandpass filter set 09 BP450-490, FT510, LP515, the filter set 02 G365, FT395, LP420 and the filter set 15 BP546, FT580, LP590 (all Carl Zeiss, Oberkochen, Germany).

Digital image processing was done with a Leica DFC 300FX fluorescence camera (Leica Microsystems, Heerbrugg, CH) and the software ImageAcces Premium Version 5 (Imagic AG, Galttbrugg, CH).

Thiazine Red R Displacement Assay

Recombinant human microtubule associated tau protein purified from *E.coli* was aggregated at a concentration of 5 μM with arachidonic acid (100 μM) in Tris 10 mM pH8, 24h at 37°C. Synthetic A β_{40} was aggregated at a concentration of 50 μM with arachidonic acid (100 μM) in Tris 10 mM pH8, for three days at 37°C, under shaking at 150 rpm.

Thiazine Red R was added at the concentration corresponding to the K_d of the respective aggregated protein binding site to induce a fluorescent signal that can be inhibited by the addition of a displacer compound. (K_d for aggregated tau = 18 nM, K_d for aggregated A β = 49 nM).

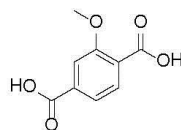
To determine the affinity of a displacer compound to the *Thiazine Red R* binding sites of the aggregated proteins, the compound was added at different concentrations to the assay ranging from 0.1 nM to 10000 nM. For the inhibition curve, the compound was measured together with the aggregated proteins and *Thiazine Red R*. The fluorescence of some dyes overlapped with the fluorescence of *Thiazine Red R* when measured at 595 nm. Therefore, compounds were also measured together with the aggregated proteins, but without *Thiazine Red R* (autofluorescence curve). Net-fluorescence was calculated by subtracting the fluorescence of the wells without *Thiazine Red R* from the fluorescence of wells containing *Thiazine Red R* (inhibition curve). As negative control *Thiazine Red R* and aggregated protein was used. As positive control, *Thiazine Red R*, reference compound with known activity (see Table 1) and aggregated protein was used.

The assay was performed in Perkin Elmer OptiPlate 384, black, 45 mL assay volume. As assay buffer DPBS (no CaCl₂ no MgCl₂) (GIBCO N. 14020) was used. The tested compounds were diluted in DMSO and 2 µL of the solution was added to the assay (5% DMSO final). The assay was started by the addition of the aggregated protein (competitive condition). The plates were shortly shaken (1 min with Sterico variomag teleshake) and incubated at room temperature for 30 min. Measurements were performed with En:Vision (Perkin Elmer), at Excitation 531 nm / Emission 595 nm.

Experimental Conditions

All chemicals were purchased as reagent grade from commercial suppliers and used without further purification, unless otherwise noted. Proton (¹H NMR) and carbon (¹³C NMR) nuclear magnetic resonance spectra were recorded on Bruker DRX 500 at 500 MHz and 125.75 MHz respectively at 293 K. The chemical shifts are given in parts per million (ppm) on the delta scale (δ) and are referenced to tetramethylsilane (δ = 0 ppm), ¹H NMR. Abbreviations for NMR data: s = singlet; d = doublet; t = triplet; q = quartet; qi = quintet; dd = doublet of doublets; dt = doublet of triplets; dq = doublet of quartets; tt = triplet of triplets; m = multiplet. Mass spectra were recorded on a Finnigan MAT 95 magnetic sector spectrometer. Thin layer chromatography (TLC) was performed using Merck TLC silica gel 60 F 254 (0.2 mm) on Aluminium sheets. Silica gel columns for chromatography were prepared with E. Merck silica gel 60 (0.063-0.20 mesh ASTM). UV/vis-spectra in methanolic solution were recorded using a Shimadzu UV-2401 PC UV-VIS Recording Spectrometer, fluorescence spectra in methanolic solution were recorded using a Shimadzu RF-1501 Spectrofluorophotometer. 2D fluorescence spectra were recorded on a Tecan Infifite® M1000. HPLC analysis was carried out using: 1) an Agilent 1100 with a reversed phase column (Zorbax Eclipse XDB-C8; 4.6*150 mm) and a 254 nm detector. The eluent is composed of H₂O (1% TFA) (A) and acetonitrile (B) with a gradient: 30 to 90%.

2-Methoxybenzene-1,4-dioic acid 4



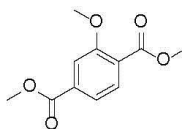
2,5 dimethylanisole (7.00 g, 46.0 mmol) was dissolved in hot water (700 ml). KMnO_4 (25.46 g, 161.0 mmol) was added in small portions over 1 h at 100°C to the vigorously stirred solution. The reaction mixture was stirred over night at 100°C . The precipitated MnO_2 was filtered off from the hot solution and washed twice with boiling water. The combined filtrates were acidified with concentrated hydrochloric acid and cooled to 7°C . The precipitated diacid was filtered off after 20 h (6.10 g, 63%).

melting point: 274.4°C

^1H NMR (500 MHz, DMSO) δ = 7.69 (d, 4J = 7.8 Hz, 1H), 7.60-7.53 (m, 2H), 3.87 (s, 3H) ppm.

^{13}C NMR (125 MHz, DMSO) δ = 166.6, 166.3, 157.2, 134.2, 130.0, 125.4, 120.6, 112.2, 55.5 ppm.

Dimethyl 2-methoxyterephthalate 5



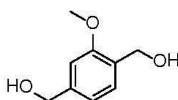
2-Methoxybenzene-1, 4-dioic acid **4** (6.00 g, 28.3 mmol) was dissolved in methanol (80 mL), cooled to 0°C and SOCl_2 was added dropwise under vigorous stirring. The reaction mixture was consecutively stirred for 30 min at 0°C , 2 h at RT and finally heated to reflux for 1 h. K_2CO_3 (2.99 g, 28.3 mmol) and Na_2CO_3 (3.90 g, 28.3 mmol) were added to the reaction mixture which was heated to reflux for 3 h. The reaction mixture was cooled to RT, quenched with ice cold water and extracted with dichloromethane. The combined organic phases were consecutively washed with 2N NaOH_{aq} and brine (three times each), dried over MgSO_4 . The solvent was evaporated in vacuum to provide 5.30 g (92%) of dimethyl 2-methoxyterephthalate **5** as colorless solid.

^1H NMR (300 MHz, DMSO) δ = 7.73 (d, 3J = 8.4 Hz, 1H), 7.59 (m, 2H), 3.88 (s, 6H), 3.81 (s, 3H) ppm.

^{13}C NMR (75 MHz, DMSO) δ = 165.3, 165.0, 157.3, 133.3, 130.4, 124.2, 120.5, 112.1, 55.6, 52.2, 51.9 ppm.

MS (m/z, 70eV, EI) = 225 (M^+), 193, 180, 149.

(2-Methoxy-1,4-phenylene)-dimethanol 6



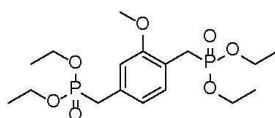
Dimethyl 2-methoxyterephthalate **5** (5.30 g, 23.6 mmol in 25 mL dry tetrahydrofuran) was added dropwise to a stirred suspension of LiAlH_4 (5.40 g, 141.9 mmol) in abs. THF (15 mL). The reaction mixture was heated to reflux for 1 h then cooled to RT. To the suspension were added slowly 6 mL of water, then 12 mL of 10 % NaOH_{aq} and again 24 mL of water. The suspension was stirred for 1 h. The aluminum salts were filtered off using a pressed plug of Celite. The combined filtrates were acidified with 2N H_2SO_4 , extracted with ethylacetate five times. The combined organic phases were washed with brine, dried over MgSO_4 and the solvent was removed in vacuo to provide 3.4 g (75%) of (2-methoxy-1, 4-phenylene) dimethanol **6**.

^1H NMR (300 MHz, CDCl_3) δ = 7.18 (d, 3J = 7.5 Hz, 1H), 6.86-6.81 (m, 2H), 4.71-4.52 (m, 4H), 3.81 (s, 3H) ppm.

^{13}C NMR (75 MHz, CDCl_3) δ = 157.5, 142.2, 128.6, 128.2, 118.8, 108.9, 65.0, 55.3 ppm.

MS (m/z, 70eV, EI) = 219 ($\text{M}^+ + \text{Na}$)

Tetraethyl (2-methoxy-1,4-phenylene)bis(methylene)diphosphonate 3



(2-methoxy-1, 4-phenylene) dimethanol **6** (3.4 g, 15.5 mmol) was suspended in dichloromethane and SOCl_2 (3.05 mL, 43.4 mmol) was added dropwise. The reaction mixture was stirred at RT over night. The reaction mixture was filtered over silica, washed with DCM twice and the solvent was removed under vacuum. The resulting dichloride was dissolved in triethylphosphite (20 mL) and stirred 16 h at 140°C . The excess of triethylphosphite was removed via distillation at 2 mbar at 150°C to provide 5.43 g (86%) of tetraethyl (2-methoxy-

1,4 -phenylene)bis(methylene)diphosphonate **3** as clear oil, slowly crystallizing as colorless solid.

¹H NMR (500 MHz, CDCl₃) δ = 7.22 (dd, ³J = 7.6 Hz, ⁴J = 2.7 Hz, 1H), 6.83-6.79 (m, 2H), 4.06-3.93 (m, 8H), 3.81 (s, 3H), 3.22 (d, ³J = 1.1 Hz, 1H), 3.15 (d, ³J = 1.1 Hz, 1H), 3.13 (d, ³J = 1.1 Hz, 1H), 3.06 (d, ³J = 1.1 Hz, 1H), 1.24-1.18 (m, 12H) ppm.

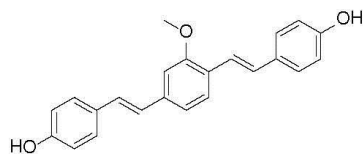
¹³C NMR (125 MHz, CDCl₃) δ = 157.2, 131.7, 131.1, 121.9, 112.9, 62.1, 55.5, 33.8 (d, J = 136.7 Hz), 26.4 (d, J = 137.8 Hz), 16.3 ppm.

MS (m/z, 70eV, EI) = 408 (M⁺ + Na), 271

Synthesis of bisphenolic distyrylbenzene-derivatives **7**

An oven-dried (140°C, 24 h) 2.5 mL Wheaton-vial was charged with KOtBu (430.5 mg; 0.83 mmol), dry DMF (1 mL), and tetraethyl (2-methoxy-1,4-phenylene)bis(methylene)-diphosphonate (200 mg, 0.48 mmol) dissolved in 0.5 mL DMF. The mixture was stirred at ambient temperature for 30 min. The corresponding aldehyde (0.83 mmol dissolved in 0.5 mL dry DMF) was added dropwise to the suspension. The reaction mixture was stirred at ambient temperature over night. The reaction mixture was transferred with 20 mL of 1N hydrochloric acid to an extraction funnel and extracted with ethyl acetate (3x 50 mL). The combined organic phases were consecutively washed with 1N hydrochloric acid, brine, dried over Na₂SO₄ and the solvent was removed under vacuum. The crude product was purified by column chromatography (hexane/ethyl acetate; 2/1) to yield the desired product.

4,4'-(1E,1'E)-2,2'-(2-methoxy-1,4-phenylene)bis(ethene-2,1-diyl)diphenol (Methoxy-XO4) **7a**

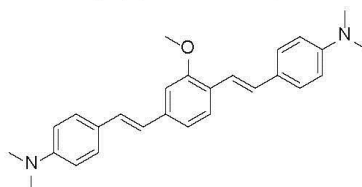


¹H-NMR (DMSO, 300 MHz): δ = 7.57 (d, ³J = 8.1 Hz, 1H), 7.42 (d, ³J = 8.6 Hz, 2H), 7.37 (d, ³J = 8.6 Hz, 2H), 7.19 (d, ³J = 16.5 Hz, 2H), 7.17 (d, ⁴J = 1.5 Hz, 1H), 7.12 (dd, ⁴J = 1.5 Hz, ³J = 8.1 Hz, 1H), 7.10 (d, ³J = 16.5 Hz, 1H), 7.00 (d, ³J = 16.5 Hz, 1H), 6.77 (d, ³J = 8.6 Hz, 2H), 6.76 (d, ³J = 8.6 Hz, 2H), 3.90 (s, 3H) ppm.

¹³C-NMR (DMSO, 75 MHz): δ = 157.2, 157.1, 156.3, 137.6, 128.6, 128.2, 127.8, 127.6, 125.8, 125.0, 124.8, 119.3, 118.8, 115.5, 108.5, 55.4 ppm.

MS (EI, 70 eV): m/z = 344 (M^+)

4-(4-(4-(Dimethylamino)styryl)-3-methoxystyryl)-*N,N*-dimethylbenzamine **7b**

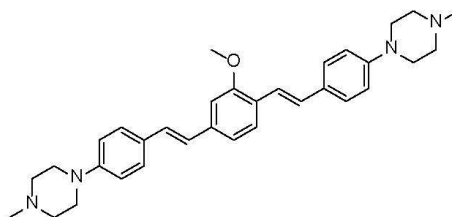


HPLC t_R = 4.26 min (90%).

¹H-NMR (DMSO- d_6 , 500 MHz): δ = 7.49 (d, 3J = 8.0 Hz, 1H), 7.40-7.37 (m, 4H), 7.25 (d, 3J = 8.8 Hz, 1H), 7.19-7.16 (m, 1H), 7.06-6.99 (m, 4H), 6.83-6.75 (m, 3H), 6.64 (d, 3J = 8.9 Hz, 1H), 3.90 (s, 3H), 2.96 (s, 12H) ppm.

MS (EI) m/z = 398.4 [M]⁺.

1-(4-(4-(4-Methylpiperazin-1-yl)styryl)-2-methoxystyryl)phenyl)-4-methylpiperazine **7c**

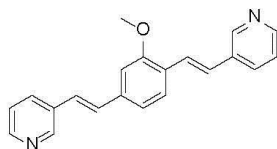


HPLC t_R = 4.65 min (83 %).

¹H-NMR (CDCl₃, 500 MHz): δ = 7.54 (d, 3J = 8.0 Hz, 1H), 7.44 (t, 3J = 8.2, 8.2 Hz, 4H), 7.32 (d, 3J = 16.5 Hz, 1H), 7.05-7.02 (m, 5H), 6.93-6.88 (m, 4H), 3.94 (s, 3H), 3.36-3.30 (m, 8H), 2.72-2.67 (m, 8H), 2.45 (s, 3H), 2.44 (s, 3H) ppm.

¹³C-NMR (CDCl₃, 125MHz) δ = 155.8, 149.3, 149.0, 136.8, 128.9, 128.0, 127.3, 127.0, 126.5, 125.5, 124.6, 124.2, 123.5, 123.3, 122.4, 122.2, 118.1, 107.4, 54.1 ppm.

3-(3-Methoxy-4-((E)-2-(pyridin-3-yl)vinyl)styryl)pyridine **7d**



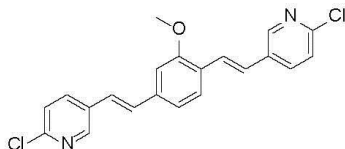
HPLC tR = 1.44 min (90 %).

$^1\text{H-NMR}$ (CDCl_3 , 500 MHz): δ = 8.76 (s, 2H), 8.51 (d, 3J = 3.8 Hz, 1H), 8.48 (d, 3J = 3.8 Hz, 1H), 7.95 (d, 3J = 8.0 Hz, 1H), 7.88 (d, 3J = 8.0 Hz, 1H), 7.60 (d, 3J = 8.0 Hz, 1H), 7.55 (d, 3J = 16.6 Hz, 1H), 7.37-7.32 (m, 2H), 7.16-7.12 (m, 5H), 3.98 (s, 3H) ppm.

$^{13}\text{C-NMR}$ (CDCl_3 , 125 MHz): δ = 155.9, 146.5, 146.4, 145.7, 145.3, 136.4, 132.7, 132.3, 131.7, 129.5, 125.5, 124.6, 124.2, 123.5, 123.3, 122.4, 122.2, 118.1, 107.4, 54.1 ppm.

MS (ESI) m/z = 314 $[\text{M}]^+$, 299.

5-(4-((E)-2-(6-Chloropyridin-3-yl)vinyl)-2-methoxystyryl)-2-chloropyridine **7e**



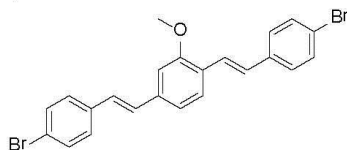
HPLC tR = 9.33 min (87 %).

$^1\text{H-NMR}$ (CDCl_3 , 500 MHz): δ = 8.43 (d, 3J = 2.4 Hz, 1H), 8.40 (d, 3J = 2.4 Hz, 1H), 7.78-7.73 (m, 2H), 7.51 (d, 3J = 8.1 Hz, 1H), 7.41 (d, 3J = 16.6 Hz, 1H), 7.26 (d, 3J = 8.3 Hz, 1H), 7.23 (d, 3J = 8.3 Hz, 1H), 7.08-7.06 (m, 2H), 7.04 (d, 3J = 2.8 Hz, 1H), 7.00 (d, 3J = 3.1 Hz, 1H), 6.96 (s, 1H), 3.90 (s, 3H) ppm.

$^{13}\text{C-NMR}$ (CDCl_3 , 125 MHz): δ = 156.3, 149.1, 148.8, 147.3, 147.2, 136.7, 134.3, 134.2, 131.6, 130.9, 130.1, 126.1, 124.9, 124.8, 123.3, 123.2, 123.1, 122.9, 118.6, 107.9, 54.6 ppm.

MS (EI) m/z = 382 $[\text{M}]^+$, 367.

1,4-Bis(4-bromostyryl)-2-methoxyneazole **7f**



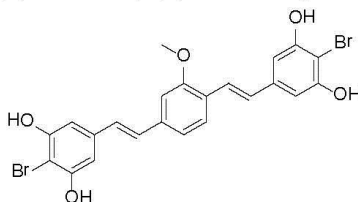
HPLC tR = 9.85 min (90 %).

$^1\text{H-NMR}$ (CDCl_3 , 500 MHz): δ = 7.48 (d, 3J = 8.0 Hz, 1H), 7.41-7.35 (m, 4H), 7.32-7.29 (m, 4H), 7.04 (d, 3J = 8.1 Hz, 2H), 7.00-6.93 (m, 4H), 3.87 (s, 3H) ppm.

$^{13}\text{C-NMR}$ (CDCl_3 , 125 MHz): δ = 157.2, 137.8, 136.9, 136.1, 131.8, 131.7, 129.2, 128.0, 127.9, 127.8, 127.5, 126.7, 125.8, 123.8, 121.4, 121.1, 119.4, 108.8, 55.6 ppm.

MS (EI) m/z = 470 (468, 472) $[M]^+$.

4-(4-(3,5-Dibromo-4-hydroxystyryl)-3-methoxystyryl)-2,6-dibromophenol 7g



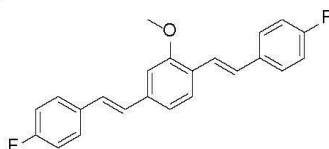
HPLC t_R = 9.85 min (90 %).

$^1\text{H-NMR}$ (DMSO- d_6 , 500 MHz): δ = 7.81 (s, 2H), 7.74 (s, 2H), 7.58 (d, 3J = 8.1 Hz, 1H), 7.28 (d, 3J = 16.5 Hz, 1H), 7.21 (d, 3J = 10.6 Hz, 3H), 7.15 (dd, 3J = 16.0, 12.6 Hz, 2H), 3.91 (s, 3H) ppm.

$^{13}\text{C-NMR}$ (DMSO- d_6 , 125 MHz): δ = 156.7, 150.0, 149.8, 137.7, 132.5, 132.0, 130.0, 129.8, 128.3, 126.7, 125.9, 125.8, 124.7, 122.9, 119.2, 112.3, 109.1, 55.5 ppm.

MS (EI) m/z = 660 $[M]^+$.

1,4-Bis(4-fluorostyryl)-2-methoxybenzene 7h



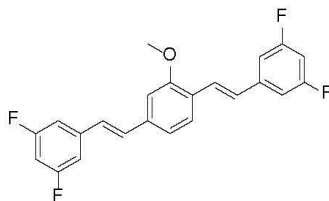
HPLC t_R = 10.13 min (80 %).

$^1\text{H-NMR}$ (CDCl_3 , 500 MHz): δ = 7.56 (d, 3J = 8.0 Hz, 1H), 7.51-7.46 (m, 4H), 7.38 (d, 3J = 16.5 Hz, 1H), 7.12-7.09 (m, 2H), 7.06-7.02 (m, 6H), 6.90-6.87 (m, 1H), 3.95 (s, 3H) ppm.

$^{13}\text{C-NMR}$ (CDCl_3 , 125 MHz): δ = 161.4, 157.1, 137.8, 134.2, 133.4, 128.3, 128.3, 128.0, 128.0, 127.8, 127.8, 127.5, 126.5, 125.9, 122.9, 119.3, 115.8, 115.6, 115.4, 114.7, 108.7, 55.6 ppm.

MS (EI) m/z = 348 $[M]^+$.

1-(4-(3,5-Difluorostyryl)-3-methoxystyryl)-3,5-difluorobenzene 7i

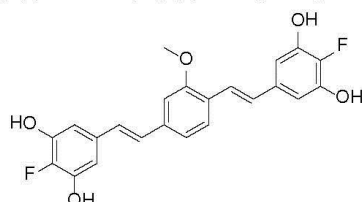


HPLC t_R = 10.28 min (90%).

$^1\text{H-NMR}$ (CDCl_3 , 500 MHz): δ = 7.48 (d, 3J = 8.0 Hz, 1H), 7.38 (d, 3J = 16.4 Hz, 1H), 7.05-7.02 (m, 2H), 6.99-6.91 (m, 7H), 6.65-6.58 (m, 2H), 3.88 (s, 3H) ppm.

^{13}C -NMR (CDCl_3 , 125 MHz): δ = 163.3, 163.2, 161.3, 161.2, 156.4, 140.4, 139.6, 136.6, 129.9, 126.1, 125.9, 124.6, 118.6, 108.2, 108.1, 108.0, 107.9, 101.9, 101.9, 101.3, 54.6 ppm.
MS (EI) m/z = 384 $[\text{M}]^+$.

4-(4-(3,5-Difluoro-4-hydroxystyryl)-3-methoxystyryl)-2,6-difluorophenole 7j



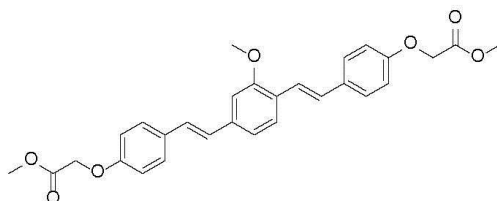
HPLC t_R = 7.93 min (81 %).

^1H -NMR (CDCl_3 , 500 MHz): δ = 7.41 (d, 3J = 8.01 Hz, 1H), 7.18 (d, 3J = 16.4 Hz, 1H), 6.99 (d, 3J = 8.2 Hz, 1H), 6.94 (d, 3J = 7.9 Hz, 4H), 6.88 (d, 3J = 8.8 Hz, 2H), 6.84 (d, 3J = 5.9 Hz, 2H), 3.85 (s, 3H) ppm.

^{13}C -NMR (CDCl_3 , 125 MHz): δ = 156.0, 152.7, 150.7, 136.5, 132.8, 132.7, 132.4, 128.1, 128.0, 127.2, 127.1, 127.0, 126.0, 125.8, 125.5, 125.1, 124.5, 121.7, 118.2, 108.4, 107.6, 54.5 ppm.

MS (EI) m/z = 416 $[\text{M}]^+$, 276.

2,2'-(4,4'-(1E,1'E)-2,2'-(2-Methoxy-1,4-phenylene)bis(ethendiyl)bis(4,1-phenylene))bis(oxy)diacetate 7k



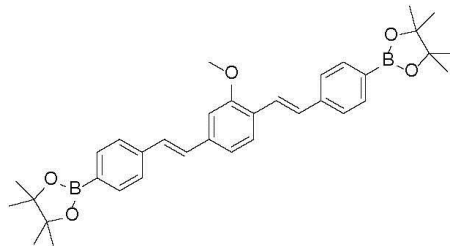
HPLC t_R = 9.05 min (63%).

^1H -NMR (CDCl_3 , 500 MHz): δ = 7.47 (d, 3J = 8.14 Hz, 1H), 7.41 (d, 3J = 8.5 Hz, 2H), 7.39 (d, 3J = 8.5 Hz, 2H), 7.28 (d, 3J = 16.4 Hz, 1H), 7.03 (d, 3J = 8.14 Hz, 1H), 7.01 (d, 3J = 16.4 Hz, 1H), 6.99 (d, 3J = 1.5 Hz, 1H), 6.93 (s, 1H), 6.90 (d, 3J = 15.9 Hz, 1H), 6.84 (d, 3J = 8.5 Hz, 2H), 6.83 (d, 3J = 8.5 Hz, 2H), 4.59 (s, 2H), 4.58 (s, 2H), 3.87 (s, 3H), 3.75 (s, 3H), 3.74 (s, 3H) ppm.

^{13}C -NMR (CDCl_3 , 125 MHz): δ = 168.4, 168.3, 168.3, 156.5, 156.4, 156.2, 155.9, 136.8, 131.0, 130.3, 127.1, 126.8, 126.7, 126.2, 125.3, 124.9, 120.8, 118.2, 113.9, 113.8, 107.6, 64.4, 64.3, 54.5, 51.3, 51.2 ppm.

MS (EI) m/z = 488 $[\text{M}]^+$.

2,2'-(4,4'-(1*E*,1'*E*)-2,2'-(2-Methoxy-1,4-phenylene)bis(ethene-2,1-diyl)bis(4,1-phenylene))bis(4,4,5,5-tetramethyl-1,3,2-dioxaborolane) **7I**



HPLC tR = 6.65 min.

¹H-NMR (CDCl₃, 500 MHz): δ = 7.80 (t, ³J = 8.80 Hz, 4H), 7.60 (d, ³J = 8.0 Hz, 1H), 7.55 (d, ³J = 8.0 Hz, 1H), 7.53 (d, ³J = 7.0 Hz, 4H), 7.17 (d, ³J = 16.3 Hz, 2H), 7.16 (d, ³J = 16.3 Hz, 2H), 7.05 (s, 1H), 3.96 (s, 3H), 1.36 (s, 24H) ppm.

¹³C-NMR (CDCl₃, 125 MHz): δ = 157.2, 140.7, 139.9, 137.9, 135.2, 135.1, 131.7, 129.5, 129.0, 128.7, 126.7, 126.1, 125.9, 125.8, 124.1, 119.5, 108.84, 83.8, 55.6, 24.9 ppm.

MS (EI) m/z = 564 [M]⁺.

3.5 Untersuchung der Struktur/Aktivitäts-Beziehung (SAR) von β -Carbolinen als DYRK1A-Inhibitoren und Steigerung der Selektivität gegenüber MAO-A mit Hilfe moderner *in silico*-Methoden.

Der Inhalt dieses Kapitels wurde bereits veröffentlicht:

Binia Drung, Christoph Scholz, Valéria A. Barbosa, Azadeh Nazari, Maria H. Sarragiotto, Boris Schmidt, "Computational & experimental evaluation of the structure/activity relationship of beta-carbolines as DYRK1A inhibitors", *Bioorganic and Medicinal Chemistry Letters* **2014**, 24, 4854-4860.

Mit freundlicher Genehmigung des Verlags Elsevier Limited. Nummerierungen wie im Manuskript.

Durch das Zusammenspiel ligandenbasierter Pharmakophor-Modellierung (siehe Abschnitt 1.2.1), strukturbasiertem Protein/Ligand-Docking (siehe Abschnitt 1.2.3), chemischer Synthese und *in vitro*-Evaluation wurde in dieser Arbeit ein äußerst interdisziplinärer Ansatz verfolgt, um möglichst viele Facetten der Struktur/Aktivitäts-Beziehung von β -Carbolinen als DYRK1A-Inhibitoren zu beleuchten. Darauf aufbauend konnten die gewonnenen Informationen genutzt werden, um die Selektivität der Leitstruktur Harmin **1** gegenüber MAO-A signifikant zu verbessern, ohne dabei an inhibitorischer Aktivität bezüglich DYRK1A zu verlieren.

Zunächst wurden auf Grundlage eines Trainings-Sets, bestehend aus 18 aktiven und 68 inaktiven Verbindungen, sieben 5-Pkt.-Ph4-Hypothesen generiert, die anschließend gegen ein Test-Set, bestehend aus 32 inaktiven und 9 aktiven Verbindungen, validiert wurden. Das Ph4-Modell, welches die effektivste Selektion aktiver und inaktiver Verbindungen des Test-Sets unter gleichzeitig möglichst guter Ausrichtung aller aktiver Verbindungen an den Pharmakophor ermöglichte, wurde identifiziert und ausgewertet. Einige essentielle Ph4-Eigenschaften sind jedoch auch durch ein noch so gutes Modell nicht hinreichend genau beschrieben. Beispielsweise kann ein hydrophober Bereich im Ph4-Modell durch eine Vielzahl chemischer Pendants erfüllt werden: Unterschiedlich lange oder verzweigte Alkylketten, Aryleinheiten, etc. Aus diesem Grund wurden Verbindungen, die genau diese Eigenschaften abtasten sollten, synthetisiert und *in vitro* auf ihre inhibitorische Aktivität bezüglich DYRK1A evaluiert. Auf diese Weise gelang es, alle Ph4-Eigenschaften auf ihre Signifikanz und eventuelle Limitierungen hin zu untersuchen und das computergenerierte Modell empirisch zu verifizieren.

Neben der Analyse der SAR war das Hauptziel dieses Projektes die Optimierung der Leitstruktur Harmin **1** bezüglich ihrer Selektivität gegenüber MAO-A ($IC_{50,DYRK1A} = 80 \text{ nM}$, $K_{i,MAO-A} = 5 \text{ nM}$). Die Ko-Inhibition dieser Oxidoreduktase führt zu unerwünschten psychoaktiven Nebenwirkungen, die den Einsatz von β -Carbolinen in klinischen Studien bislang verhindern. Durch Protein/Ligand-Docking und Analyse der wissensbasierten statistischen Potentiale der *Scoring*-Funktion DSX wurden drei Derivate als besonders vielversprechend identifiziert und deren inhibitorische Aktivität gegenüber MAO-A *in vitro* bestimmt. Das N-Heptyl-Derivat **16** inhibierte MAO-A bei einer Konzentration von $1 \mu\text{M}$ nur zu 13% (vgl. Harmin **1** = 103%).

Aufgrund des synergistischen Zusammenspiels verschiedenster Methoden des modernen Wirkstoffdesigns konnte diese deutliche Selektivitätssteigerung bei ähnlich guter inhibitorischer Aktivität des N-Heptyl-Derivats **16** ($IC_{50,DYRK1A} = 130 \text{ nM}$) erreicht werden und bereitet β -Carbolin-Derivaten möglicherweise den Weg in erste klinische Studien als DYRK1A-Inhibitoren.

Beitrag C. Scholz: Verfassen des Manuskriptes (zusammen mit B. Drung), Auswahl der Liganden-Datenbanken, Ph4-Modellierung, Analyse der SAR, Protein/Ligand-Docking.



Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl



Computational & experimental evaluation of the structure/activity relationship of β -carbolines as DYRK1A inhibitors



Binia Drung^{a,†}, Christoph Scholz^{a,†}, Valéria A. Barbosa^b, Azadeh Nazari^a, Maria H. Sarrajiotto^b, Boris Schmidt^{a,*}

^aClemens Schöpf–Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Alarich-Weiss-Str. 4, 64287 Darmstadt, Germany

^bDepartamento de Química, Universidade Estadual de Maringá, Avenida Colombo 53790, PR 87020-900 Maringá, Brazil

ARTICLE INFO

Article history:

Received 25 June 2014

Revised 22 August 2014

Accepted 25 August 2014

Available online 2 September 2014

Keywords:

DYRK1A

MAO-A

Pharmacophore

Docking

Inhibitor

β -Carboline

SAR

ABSTRACT

DYRK1A has been associated with Down's syndrome and neurodegenerative diseases, therefore it is an important target for novel pharmacological interventions. We combined a ligand-based pharmacophore design with a structure-based protein/ligand docking using the software MOE in order to evaluate the underlying structure/activity relationship. Based on this knowledge we synthesized several novel β -carboline derivatives to validate the theoretical model. Furthermore we identified a modified lead structure as a potent DYRK1A inhibitor (IC_{50} = 130 nM) with significant selectivity against MAO-A, DYRK2, DYRK3, DYRK4 & CLK2.

© 2014 Elsevier Ltd. All rights reserved.

The dual specificity tyrosine phosphorylation regulated kinase-1A (DYRK1A) is one of five members of a eukaryotic kinase family. These kinases are named according to their capability to (1) self-activate by autophosphorylation of Tyr321, which is located in the so called activation loop, and to (2) phosphorylate serine and threonine residues on target proteins.¹ The gene for DYRK1A is the only one of the five subtypes that is located on the human chromosome 21 in the Down's syndrome critical region.² It is linked to Down's syndrome and neurodegenerative diseases such as Alzheimer's disease, Huntington's disease, Parkinson's disease and Pick's disease.³ The demographic changes affecting the western societies place more people at risk. For this reason rising healthcare costs seem unavoidable. New drug targets for neurodegenerative diseases are thus in strong demand and DYRK1A may be an important target for pharmacological intervention.

Several DYRK1A inhibitors displaying high activity and moderate kinase selectivity have been identified by compound screening in kinase assay panels. Despite these efforts, harmine **1** is still the most active and orally bioavailable DYRK1A inhibitor. Yet it holds

very limited therapeutic potential as it generates several side effects which rule out chronic clinical application.² Both **1** and the closely related harmol **2** are potent MAO-A (monoamine oxidase A) inhibitors void of psychoactive properties, which imposes a safety alert and clearly limits their use in animals.^{1,4} The MAO-A inhibition can result in extended half-lives of psychoactive drugs and thus **1** and **2** are employed in psychoactive drug formulations.⁵ Both belong to the β -carboline alkaloids, which occur in several plants. The most important natural sources are the Syrian rue (*Peganum harmala*) and a South American vine (*Banisteriopsis caapi*). In the Amazon region this vine is used traditionally by shamans to prepare a brew, called ayahuasca, which induces psychoactive effects.⁶ Pharmacological studies of the brew and especially **1** revealed an inhibitory effect of MAO-A. Besides **1**, several other β -carboline alkaloids have been screened and found to inhibit DYRK1A.⁷ Compound **1** has been reported to selectively inhibit DYRK1A in a panel of 69 kinases. Treating the kinase with 1 μ M concentration of the inhibitor caused a remaining activity of $4 \pm 2\%$ (IC_{50} = 80 nM). In this panel of protein kinases the two family members DYRK2 ($15 \pm 0\%$, IC_{50} = 900 nM) and DYRK3 ($6 \pm 0\%$, IC_{50} = 800 nM) were inhibited to a lesser extent.^{7b} Depending on the assay's conditions, particularly the concentration of ATP, the IC_{50} of DYRK1A inhibition by **1** varied in a range from 33 nM to 700 nM.^{2,7,8}

* Corresponding author. Tel.: +49 6151 163075.

E-mail address: schmidt_boris@t-online.de (B. Schmidt).

[†] These authors contributed equally. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

For a rational design of new DYRK1A inhibitors several computational methods are available, which can be grouped into two major classes: ligand- and structure-based molecular modeling.⁹ The latter is enabled by an X-ray structure elucidation of the **1**/DYRK1A complex reported by Ogawa et al. in 2010.¹⁰ In this work we utilized the software package MOE¹¹ to combine a ligand-based pharmacophore design together with a structure-based protein/ligand docking to (1) estimate essential ligand/receptor interactions across a set of inhibitor classes, (2) distinguish active from inactive compounds and (3) postulate suitable binding modes of our newly synthesized inhibitors without the need of additional X-ray crystallographic analyses. The technique of pharmacophore elucidation is based on the hypothesis that even structurally and chemically diverse compounds share certain features on specific positions like hydrophobic regions, aromatic rings and H-bond donors or acceptors, provided that they target the same active site of a protein. In a recent work Pan et al. disclosed the first pharmacophore model for DYRK1A exclusively on the basis of 6-arylquinazolin-4-amines which is presumably not adaptive to other inhibitor classes.¹² Our intent was to broaden the applicability of the pharmacophore model by evaluating a consistent pharmacophore among a set of 127 structurally diverse known active ($IC_{50} \leq 250$ nM) or inactive ($IC_{50} \geq 10$ μ M) compounds from 16 different publications.¹³ Especially for the development of pharmacophore-based virtual screenings, the diversity of the training set is of importance to avoid a possible over-fitting of the model, which could lead to a high amount of false-negative rated compounds. To validate our modeling approach in terms of structural significance and limitations, an in-house test set of new compounds was synthesized and these were evaluated for their inhibitory activities against DYRK1A.

The pharmacophore models were generated with 18 active compounds of the training set and validated against 68 inactives. The elucidation resulted in a total of seven different five-feature hypotheses, with four of them leading to an overall accuracy $\geq 80\%$ (see [Supplementary material](#) for details).¹⁴ These four pharmacophores were evaluated against 32 inactive and 9 active test set compounds. All four pharmacophores showed a high overall accuracy of 90% in the evaluation step, covering 9 of 9 actives and only 4 of 32 inactives. To rate the models despite of their identical accuracy values, the root of the mean square distances (RMSD) between the pharmacophore features and the ligand annotation points of all active test set compounds were taken into account. [Figure 1](#) shows **1** aligned to the highest rated pharmacophore model. It consists of two aromatic features (F1 & F2), fulfilled by the indole scaffold of **1**, two hydrophobic centroids (F3 & F4) covered by a methyl and a methoxy substituent at the exterior of the ligand and one H-bond acceptor projection (F5) pointing towards the pyridyl nitrogen. This projection is synonymous to

an H-bond donor atom of the receptor (Lys188) whose identification is possible for example by protein/ligand docking.

To validate that the docking procedure is able to predict correct binding modes of DYRK1A inhibitors, a redock of **1** into the X-ray structure of the **1**/DYRK1A complex was performed ([Fig. 2](#)). The DSX scoring function impressively demonstrated its ability to identify near native binding modes (see [Supplementary material](#) for details).¹⁵ The two docking poses with the right overall orientation compared to the reference ligand were assigned with the highest scores and rated in correct order among each other. According to the London dG score, these poses are rated at the fourth and fifth position and in reverse order regarding to their RMSD.

To validate the pharmacophore model in matters of structural significance and limitations, an in-house test set of compounds was synthesized. Since **1** is one of the most promising lead structures for DYRK1A inhibitors, the scaffold was selected for derivatization.² The in-house test set consisted of 16 potential DYRK1A inhibitors which were aligned to the developed pharmacophore model. The alignment was successful (RMSD < 0.8 Å) for 10 compounds. No compound rated as inactive (inhibition $\leq 50\%$ at 10 μ M) in the later biological testing was found in this set of successfully aligned molecules ([Tables 1 and 3](#)). In the set of the six molecules that failed the pharmacophore search, three active substances were identified which exerted an inhibition $> 50\%$ and are therefore classified as false-negatives.

The experimental determination of the inhibitory activity for DYRK1A of the synthesized compounds was performed by Cerep according to a published procedure based on a FRET between labeled antibodies and the substrate.¹⁶ For reference, Staurosporine **3** was used.² Compounds **1** and **2** were tested under the same conditions. The results of the in vitro assay are expressed as percentage of control specific activity obtained in the presence of 10 μ M of the test compound ([Table 1](#)).

Apparently a hydrophobic region, as covered by pharmacophore feature 3, is desirable for all active derivatives ([Fig. 1](#)). As this constraint is already fulfilled by a methyl group and no information about the limitations of this hydrophobic sphere is presented in the pharmacophore model, an extension of the methyl moiety to ethyl and *n*-propyl was tested. Therefore 6-methoxytryptophane **4** and the corresponding aldehydes were combined in a Pictet–Spengler cyclization to form the intermediate **5**.¹⁷ Oxidation of **5** with potassium dichromate yielded in ethyl and *n*-propyl elongated carboline derivatives **6** and **7** ([Scheme 1](#)).

The percentage of inhibition decreased with an elongation of methyl to ethyl and *n*-propyl. **1** exerted an inhibition of 97.3% at a concentration of 10 μ M. **6** containing an ethyl moiety inhibited DYRK1A at 96.7% and **7** with an *n*-propyl sidechain inhibited 96.1% of DYRK1A activity. This trend becomes even more apparent with the additional measurement at a concentration of 1 μ M (see [Table 2](#)).

In parallel three β -carboline derivatives void of the 7-methoxy group were synthesized. The elongation series contained harmine **8**, bearing a methyl group featuring the smallest qualification for the pharmacophore model, and two extended derivatives with an ethyl **9** and *iso*-butyl moiety **10**. The synthesis employed tryptophane **11** as starting material with acetaldehyde, propionaldehyde or methylbutyraldehyde, following the same reaction procedure as for **6** and **7**.

Compound **8**, the lead structure of the second series without a methoxy substituent, displayed a reduced inhibitory activity (80.3%) compared to **1**. An extension to ethyl in **9** reduced the inhibitory activity to 49.6% and the bulkier *iso*-butyl in **10** resulted in a drop of activity to 32.8%. Both series of chain extensions, covered by the hydrophobic constraint F3, are characterized by the same structure/activity trend. In conclusion, these results show that any alkyl chain longer than methyl decreases the inhibitory

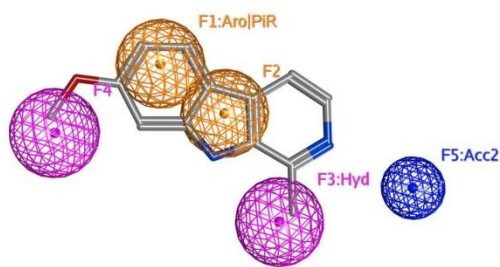


Figure 1. The highest rated pharmacophore model mapped to **1**. Purple (F3: Hyd, F4: Hydrophobic centroid, orange (F1: Aro/PIR, F2: Aromatic system, blue (F5: Acc2): H-bond acceptor projection.

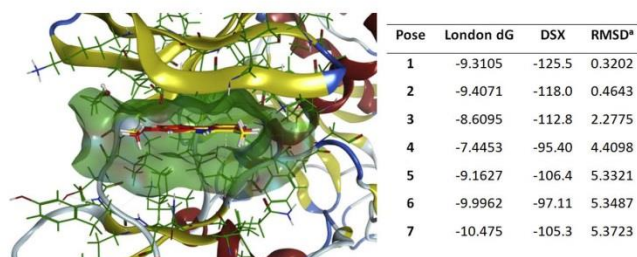


Figure 2. Redock of **1** in the X-ray structure of the **1**/DYRK1A complex, the corresponding root of the mean square distances (RMSD) and score values. ^aRMSD from reference ligand in Å. Green: Protein carbon atoms, yellow: Reference ligand from PDB-ID: 1ANR, red: Redocked pose with the highest DSX score (RMSD = 0.3202 Å). DSX: drug score extended.¹⁵

Table 1
Determination of % inhibition for DYRK1A at 10 μ M

Compound	R ¹	R ²	R ³	R ⁴	RMSD ^a	% Inhibition at 10 μ M ^b
1	OMe	H	Me	H	0.4791	97.3 \pm 1.1
2	OH	H	Me	H	N/A	99.3 \pm 1.1
6	OMe	H	Et	H	0.5004	96.7 \pm 0.1
7	OMe	H	<i>n</i> -Pr	H	0.4906	96.1 \pm 0.3
8	H	H	Me	H	N/A	80.3 \pm 0.3
9	H	H	Et	H	N/A	49.6 \pm 1.9
10	H	H	<i>i</i> -Bu	H	N/A	32.8 \pm 3.0
12	OEt	H	Me	H	0.7349	92.4 \pm 0.6
13	OHep	H	Me	H	0.5107	84.2 \pm 0.4
14	(OCH ₂ CH ₂) ₂ OMe	H	Me	H	0.5252	90.9 \pm 1.1
15	(OCH ₂ CH ₂) ₃ OMe	H	Me	H	0.5694	88.9 \pm 1.1
16^c	OMe	Hep	Me	H	0.4963	97.6 \pm 1.7
17	OMe	H	Me	6-Cl	0.4660	95.0 \pm 1.1
18	OMe	H	Me	8-Cl	N/A	91.2 \pm 0.2

^a Root of the mean square distances (RMSD) of pharmacophore features to their respective ligand annotation points in Å. Compounds that failed the pharmacophore alignment are annotated with N/A.

^b Results are expressed as a percentage of control specific activity = 100 – ((measured specific activity/control specific activity)*100) obtained in the presence of 10 μ M of the test compounds. The measurements were performed in duplicates.

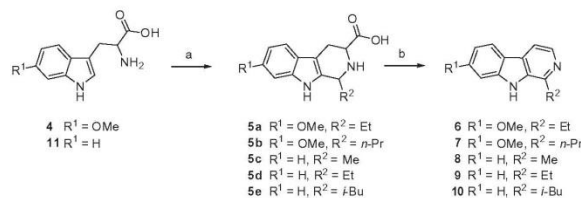
^c IC₅₀ = 130 nM.

activity, while the hydrophobic centroid is an essential constraint according to the pharmacophore model. Since the protein structure of DYRK1A and the binding mode of **1** within the protein is known from X-ray analysis, an important H-bond interaction of the methoxy oxygen and the backbone amide of Leu241 is claimed.^{1,10} This effect is apparent in the decrease of inhibitory activity between **1** and **8**.

Next we explored the limitations of the hydrophobic feature F4 by an elongation analysis. Therefore a set of compounds containing an ethoxy **12**, a heptyloxy **13** and also pegylated sidechains in two different lengths **14** and **15**, were synthesized. The ether functionality was retained to utilize the interactions of the methoxy moiety and Leu241. Compound **2** as starting material was alkylated with the corresponding tosylated or halogenated reagents in the presence of cesium carbonate. The enlargement of the ether groups resulted in only a small decrease in inhibitory activity. **12** still inhibits DYRK1A at 92.4% and the elongation to **13** decreased the inhibitory activity to 84.2%.

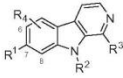
The well-established solubilizing effect of polyethyleneglycol chains (PEG's) encouraged us to switch the heptylether chain in **13** to a diethyleneglycol chain of the same length as realized in compound **14**.¹⁸ This resulted in an increase of inhibitory activity back to 90.9% for **14**. Addition of one ethylene group to a tripegylated chain in **15** lowered the activity by a non-significant amount to 88.9%. Hence all four derivatives are potent DYRK1A inhibitors, the methoxy position in **1** seems quite variable for derivatisations. This result could also be explained by protein/ligand docking, as the alkyl and PEG-chains point out of the active site and therefore the pharmacophore feature F4 is covered by a variety of flexible hydrophobic chains (Fig. 3).

No information about substituents at the N-H-moiety of **1** was provided by the pharmacophore model. The structure based protein/ligand docking predicted a free space in this position. The interaction between **1** and DYRK1A aligns the N-H-position towards a solvent exposed region of the active site, where no steric hindrance is expected (see Fig. S1 in the Supplementary material). **16** containing a heptyl chain in the former N-H-position was synthesized by reaction of **1** with iodoheptane in presence of sodium hydride. The inhibition assay revealed an inhibitory activity of 97.6% for **16** and thus confirmed the computational results. In the selectivity study presented at a later stage in this work **16** was identified as the most promising compound, therefore the



Scheme 1. Reagents and conditions. (a) R²COH, AcOH, 70 °C, 2–3 h; (b) K₂Cr₂O₇(aq), AcOH, reflux, 3–6 h.

Table 2
Determination of % inhibition for DYRK1A at 1 μ M



Compound	R ¹	R ²	R ³	R ⁴	% Inhibition at 1 μ M ^a
1	OMe	H	Me	H	94.0 \pm 0.1
6	OMe	H	Et	H	76.9 \pm 0.7
7	OMe	H	n-Pr	H	74.7 \pm 2.1
16	OMe	Hep	Me	H	85.0 \pm 0.5

^a Results are expressed as a percentage of control specific activity = 100 – ((measured specific activity/control specific activity)*100) obtained in the presence of 1 μ M of the test compounds. The measurements were performed in duplicates.

IC₅₀ = 130 nM was determined in addition to the single point analysis.¹⁶

In order to examine the influence of electron density in the aromatic system on the biological activity, two chloro harmine derivatives **17** and **18** were prepared. The synthesis followed a procedure by Ponce et al. to provide **17** with a chlorine atom in position 6 and **18** with a chlorine atom in position 8.¹⁹ An alignment of both compounds to the pharmacophore model revealed, that only the 6-chloro-harmine **17** passes the requirements (Table 1). The analysis of the protein/ligand docking indicated an additional interaction of the chlorine atom in position 6 to the backbone of Glu239 (see Fig. S2 in the Supplementary material). The biological data confirmed the predictions, with an inhibitory activity of 95.0% for **17** and 91.2% for the less favorable 8-chloro-harmine **18**. Nevertheless both compounds **17** and **18** are still potent DYRK1A inhibitors.

The importance of an acceptor atom inside or close to the aromatic system is illustrated by the Acc2-feature (F5) in the pharmacophore model. In case of the scaffold of **1**, analysis of the X-ray structure revealed that the nitrogen atom of the pyridine ring represents this acceptor and interacts with the sidechain of Lys188. This suggests that Lys188 corresponds to the Acc2 feature (F5) in the pharmacophore model.

However, the pharmacophore model did not specify the kind or exact location of the acceptor atom. For that reason a triazinoindole scaffold was composed, which contains acceptor atoms in all positions close to the original position in **1** (Scheme 2).

The synthesis was performed similar to the published procedure by Ashry et al. using 6-methoxyindole **19**, which was synthesized following the procedure reported by Mason et al.²⁰ After alkylation of the indole's amino position, using 1-bromo-3-methylbutane and potassium carbonate, intermediate **20** was



Scheme 2. Alternations of the acceptor atom position. (A) Possible acceptor atom positions.

obtained. The 1,2,4-triazino[5,6 b]indole-3-thiol scaffold **21** was built up by cyclization with thiosemicarbazide. The cyclization was performed under microwave irradiation in two steps. Compound **21** was methylated with iodomethane in presence of potassium carbonate, to give the 7-methoxyderivative **22** (Scheme 3).

The modified acceptor positions in the aromatic system in compound **22** reduced the DYRK1A inhibitory activity to 67.4% (Table 3). Apparently the position of the acceptor atom in the initial **1** provides a better inhibitory effect.

Compound **22** contains a methoxy group in the analogue position of **1**. The 8-methoxy-triazino-indole derivative **23** was synthesized in order to evaluate the positioning of the methoxy group. Therefore the synthesis for **22** was modified and started with the commercial available 5-methoxyindole **24**. Compound **23** with the shifted methoxy position was obtained after the final methylation. Switching of the methoxy position resulted in a dramatic collapse of the inhibition potential to 5.6%, which is indicative of a severe penalty in the binding process. For more information about reaction details and substance characterization consult the Supplementary material.

For the development of feasible β -carboline derived DYRK1A inhibitors it is necessary to enhance their selectivity. The main problem to be solved is the co-inhibition of MAO-A by **1**, which generates severe side effects upon treatment.¹ The protein structure and the detailed binding mode of **1** to the active site of MAO-A was published by Tsukihara et al. in 2008 and allows the structure-based enhancement of selectivity (PDB-ID: 2Z5X).²¹ The redocking of **1** into the X-ray structure of the 1/MAO-A complex resulted in an RMSD of 0.3237 Å.

In contrast to DYRK1A, where **1** binds at the exterior of the protein, the active site of MAO-A is almost completely closed (Fig. 4), due to an induced fit after inhibitor binding.²¹ However, protein/ligand docking of the O-tripegylated harmine derivative **15** revealed a narrow channel surrounded by Leu97, Phe108 and Ala111. This finding is in accordance with previous work by Wouters et al. in which the potency of MAO-A inhibition was increased by alkylation of this site, which is not favorable for DYRK1A-inhibitors.⁴

Besides the elongation of the ether moiety (R¹), N-alkylation (R²), C-alkylation (R³) and chlorination (R⁴) on position 6 and 8 were performed in this study, potentially providing several criteria

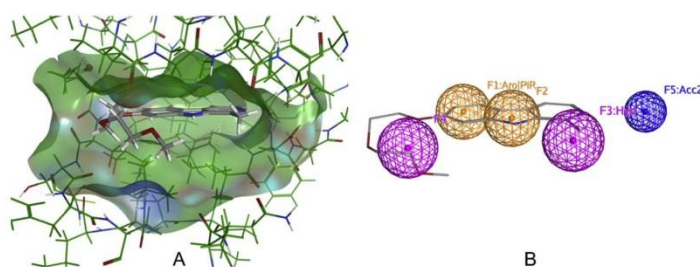
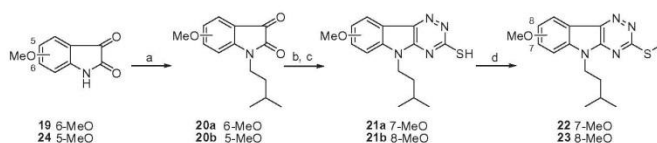
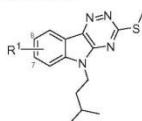


Figure 3. (A) Estimated binding mode of **14** obtained by protein/ligand docking. (B) Alignment of **14** to the pharmacophore query. Purple (F3: Hyd, F4): Hydrophobic centroid, orange (F1: Aro/PiR, F2): Aromatic system, blue (F5: Acc2): H-bond acceptor projection.



Scheme 3. Reagents and conditions. (a) $(\text{CH}_3)_2\text{CH}_2\text{CH}_2\text{Br}$, K_2CO_3 , DMF, rt; (b) $\text{NH}_2\text{CSNHNH}_2$, AcOH, EtOH, Microwave, 80 °C, 2 h; (c) K_2CO_3 , H_2O , Microwave, 100 °C, 2.5 h; (d) CH_3I , K_2CO_3 , DMF, rt.

Table 3
Determination of % inhibition for DYRK1A



Compound	R ¹	RMSD ^a	% Inhibition at 10 μM ^b
22	7-OMe	0.5813	67.4 \pm 7.0
23	8-OMe	0.8419	5.6 \pm 15.1

^a Root of the mean square distances (RMSD) of pharmacophore features to their respective ligand annotation points in Å. Compounds that failed the pharmacophore alignment are annotated with N/A.

^b Results are expressed as a percentage of control specific activity = $100 - ((\text{measured specific activity}/\text{control specific activity}) \times 100)$ obtained in the presence of 10 μM of the test compounds. The measurements were performed in duplicates.

of selectivity (Table 1). The possibility to visualize the calculated per-atom scores as spheres of different shapes is a unique feature of the DSX scoring function and enables an intuitive way to investigate a docking result (Fig. 5). Elongation of the methyl group to ethyl in **6** (Fig. 5A), *n*-propyl in **7** (Fig. 5B), and *iso*-butyl in **10** (Fig. 5C) showed a discriminating effect only with the very bulky *iso*-butyl moiety. A major decrease of the co-inhibition of MAO-A is expected for *N*-alkylation as represented in the *N*-heptyl derivative **16**. Whereas the heptyl chain points to the solvent exposed

exterior of DYRK1A and a potent inhibitory activity was observed, even the most favorable docking pose in MAO-A results in major steric clashes with the receptor (Fig. 5D). Chlorination in position 6 (derivative **17**; Fig. 5E) and in position 8 (derivative **18**; Fig. 5F) also leads to an unfavorable per atom score for MAO-A while both halogenations are well tolerated by DYRK1A.

To validate the computer-based results experimentally, we determined the inhibitory activity for MAO-A of the derivatives **7**, **16** and **17**, from which a reduced effect on MAO-A was expected. The determination was performed by Cerep using a published procedure based on a photometric analysis of 4-hydroxyquinoline.¹⁶ Compound **1** was used as an internal control compound and evaluated under the same conditions. The results of the *in vitro* assay are expressed as percentage of control specific activity obtained in the presence of 1 μM /10 μM of the test compound (Table 4).

The biological results display a complete inhibition of MAO-A by **1** (103.4%) at 1 μM , which underlines the urgent need in selectivity enhancement of DYRK1A inhibitors. All other synthesized and tested derivatives inhibit MAO-A to a significantly smaller amount. The *n*-propyl derivative **7** effects MAO-A only by 18.5%. The *N*-heptyl derivative **16**, for which the docking method predicts a major steric hindrance, possesses an even more decreased activity of 12.7% for MAO-A at 1 μM . An additional measurement at 10 μM was conducted which resulted in an inhibition of 34%. Thus the IC_{50} of **16** versus MAO-A is above 10 μM and **16** can be classified as inactive at the IC_{50} versus DYRK1A (130 nM). Only the 6-chloro derivative **17** inhibits MAO-A in a medium range of 79.9%. This could be explained by a twisted orientation of the β -carboline structure, differing from the computer-based model. A localization of the 6-chloroatom of **17** in the direction of the aforementioned

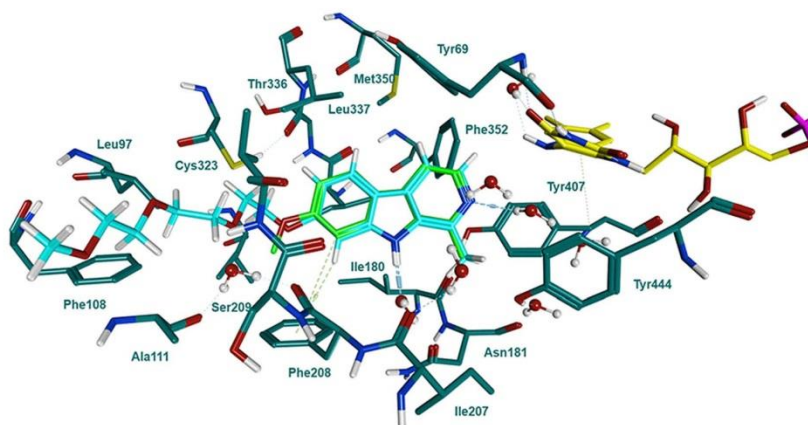


Figure 4. Binding mode of **1** (green) to the active site of MAO-A (PDB-ID: 2Z5X) and estimated binding mode of **15** (turquoise) obtained by protein/ligand docking. The essential cofactor FAD is shown in yellow.

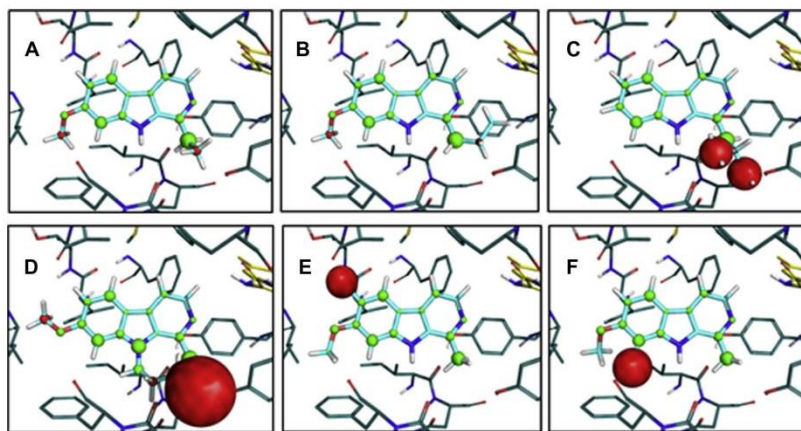
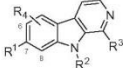


Figure 5. Graphical illustration of the calculated DSX potentials for each ligand atom. Favorable potentials are shown in green spheres whereas bad potentials are shown in red spheres; the size of a sphere represents the amplitude of the respective potential. (A) Ethyl derivative **6**, (B) *n*-propyl derivative **7**, (C) isobutyl derivative **10**, (D) *N*-heptyl derivative **16**, (E) 6-chloro derivative **17** and (F) 8-chloro derivative **18**.

Table 4
Determination of % inhibition of MAO-A

Compound					% Inhibition ^a
	R ¹	R ²	R ³	R ⁴	
1	OMe	H	Me	H	103.4 ± 0.0 (1 μM)
7	OMe	H	<i>n</i> -Pr	H	18.5 ± 0.0 (1 μM)
16	OMe	Hep	Me	H	12.7 ± 0.4 (1 μM)
16	OMe	Hep	Me	H	34.1 ± 0.4 (10 μM)
17	OMe	H	Me	6-Cl	79.9 ± 3.3 (1 μM)

^a The results are expressed as a percentage of control specific activity = 100 – ((measured specific activity/control specific activity)*100) obtained in the presence of the mentioned concentrations (1 μM/10 μM) of the test compounds. The measurements were performed in duplicates.

tunnel may circumvent the sterical clash that is responsible for the calculated unfavorable potential.

In addition we evaluated the selectivity of our most promising derivative **16** for the co-inhibition of several closely related kinases: DYRK2, DYRK3, DYRK4, CLK1 & CLK2. The measurements were performed by Cerep using published procedures.¹⁶ The results are displayed in Table 5. Compound **16** shows high selectivity against all tested kinases of the DYRK family and CLK2 (CDC2-like kinase 2) with inhibitions <30%. The co-inhibition of CLK1 at 71.0% is seen uncritical for further clinical studies since CLK1 is also a potential and well-studied target for AD and other neurodegenerative disorders.²² In fact the lead structure **1** co-inhibits CLK1 and DYRK1A equally.²³ As the ATP binding pockets of DYRK1A and CLK1 show a sequence identity of more than 70%, co-inhibition is very common.²⁴

The characterization of essential protein/ligand interactions of DYRK1A with a broad spectrum of inhibitors through a ligand-based pharmacophore elucidation will assist future lead structure identification and optimization attempts. Furthermore we validated our theoretical model by synthesis and biological evaluation of several potential DYRK1A inhibitors. The developed pharmacophore model is able to differentiate between active and inactive compounds with high accuracy, which was challenged by an exter-

Table 5
Determination of % inhibition of compound **16** for several kinases

Kinase	% Inhibition at 1 μM ^a
DYRK1A	85.0 ± 0.5
DYRK2	6.4 ± 0.5
DYRK3	28.0 ± 5.1
DYRK4	14.6 ± 0.3
CLK1	71.0 ± 3.7
CLK2	14.6 ± 0.5

^a The results are expressed as a percentage of control specific activity = 100 – ((measured specific activity/control specific activity)*100) obtained in the presence of 1 μM of compound **16**. The measurements were performed in duplicates.

nal and internal test set covering a variety of structural motifs. We identified two essential hydrophobic regions, one restricted to the size of a methyl group and one flexible for derivatisation even with steric demanding alkyl or solubilizing polyethyleneglycole chains. We further developed a step-by-step docking workflow consisting of a pharmacophore-based placement, forcefield-based pose refinement and external scoring by the knowledge-based scoring function DSX. This methodology was validated by a redocking of **1** into the published X-ray structure of the **1**/DYRK1A complex with an RMSD of only 0.3202 Å compared to the native binding mode. The N-H-position of the indole scaffold of **1** turned out to be very variable according to the protein/ligand docking, as substituents point to the solvent exposed region of the active site. Possible position variations of the essential H-bond acceptor feature, covered by the pyridine moiety of **1**, were investigated and considered unfavorable by synthesis and biological testing of triazoloindole derivatives. The pharmacophore model together with its ligand-based evaluation and in combination with reliable binding modes obtained by protein/ligand docking provide valuable knowledge for a rational design of new DYRK1A inhibitors.

For selectivity enhancement of the novel DYRK1A inhibitors, the docking procedure was also performed with the protein structure of the **1**/MAO-A complex with an RMSD of 0.3237 Å in comparison to the native binding mode. To decrease unwanted side effects induced by co-inhibition of MAO-A by **1**-based derivatives, the selectivity versus MAO-A of the novel DYRK1A inhibitors was

investigated. The four candidates with unfavorable calculated DSX potentials against MAO-A were biologically evaluated. The *N*-heptyl derivative **16** (IC₅₀(DYRK1A) = 130 nM) and the *n*-propyl derivative **7** were identified as potent DYRK1A inhibitors with significant selectivity against MAO-A. This gives rise to less side effects and may clear the way for the promising class of β -carboline-derived DYRK1A inhibitors to clinical trials.

Acknowledgments

We thank CNPq-Brazil for fellowship (Barbosa, V. A.). We thank the Hans and Ilse Breuerstiftung for support.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmcl.2014.08.054>.

References and notes

- Becker, W.; Sippl, W. *FEBS J.* **2011**, *278*, 246.
- Smith, B.; Medda, F.; Gokhale, V.; Dunckley, T.; Hulme, C. *ACS Chem. Neurosci.* **2012**, *3*, 857.
- (a) Park, J.; Song, W.-J.; Chung, K. *Cell. Mol. Life Sci.* **2009**, *66*, 3235; (b) Ferrer, I.; Barrachina, M.; Puig, B.; Martínez de Lagrán, M.; Martí, E.; Avila, J.; Dierssen, M. *Neurobiol. Dis.* **2005**, *20*, 392.
- Reniers, J.; Robert, S.; Frederick, R.; Masereel, B.; Vincent, S.; Wouters, J. *Bioorg. Med. Chem.* **2011**, *19*, 134.
- (a) Herraiz, T.; González, D.; Ancín-Azpilicueta, C.; Arán, V. J.; Guillén, H. *Food Chem. Toxicol.* **2010**, *48*, 839; (b) McIsaac, W. M.; Estevez, V. *Biochem. Pharmacol.* **1966**, *15*, 1625.
- McKenna, D. J. *Pharmacol. Ther.* **2004**, *102*, 111.
- (a) Frost, D.; Meechooet, B.; Wang, T.; Gately, S.; Giorgetti, M.; Shcherbakova, I.; Dunckley, T. *PLoS One* **2011**, *6*, e19264; (b) Bain, J.; Plater, L.; Elliott, M.; Shpiro, N.; Hastie, C. J.; McLauchlan, H.; Klevernic, I.; Arthur, J. S. C.; Alessi, D. R.; Cohen, P. *Biochem. J.* **2007**, *408*, 297.
- Göckler, N.; Jofre, G.; Papadopoulos, C.; Soppa, U.; Tejedor, F. J.; Becker, W. *FEBS J.* **2009**, *276*, 6324.
- Sheppard, A.; In *Molecular modelling: principles and applications*, 2nd ed.; Andrew, R. Leach, Ed.; Prentice Hall: 2001; p 24. ISBN-10: 0582382106.
- Ogawa, Y.; Nonaka, Y.; Goto, T.; Ohnishi, E.; Hiramatsu, T.; Kii, I.; Yoshida, M.; Ikura, T.; Onogi, H.; Shibuya, H.; Hosoya, T.; Ito, N.; Hagiwara, M. *Nat. Commun.* **2010**, *1*, 86.
- Molecular Operating Environment (MOE)*, 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7: 2013.
- Pan, Y.; Wang, Y.; Bryant, S. H. *J. Chem. Inf. Model.* **2013**, *53*, 938.
- (a) Bamford, M. J.; Bailey, N.; Davies, S.; Dean, D. K.; Francis, L.; Panchal, T. A.; Parr, C. A.; Sehmi, S.; Steadman, J. G.; Takle, A. K.; Townsend, J. T.; Wilson, D. M. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3407; (b) Zagotto, G.; Redaelli, M.; Pasquale, R.; D'Avella, D.; Cozza, G.; Denaro, L.; Pizzato, F.; Mucignat-Caretta, C. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2079; (c) Niu, X.; Dahse, H.-M.; Menzel, K.-D.; Lozach, O.; Walther, G.; Meijer, L.; Grabley, S.; Sattler, I. *J. Nat. Prod.* **2008**, *71*, 689; (d) Tang, L.; Li, M.-H.; Cao, P.; Wang, F.; Chang, W.-R.; Bach, S.; Reinhardt, J.; Ferandin, Y.; Galons, H.; Wan, Y.; Gray, N.; Meijer, L.; Jiang, T.; Liang, D.-C. *J. Biol. Chem.* **2005**, *280*, 31220; (e) Wang, D.; Wang, F.; Tan, Y.; Dong, L.; Chen, L.; Zhu, W.; Wang, H. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 168; (f) Mott, B. T.; Tanega, C.; Shen, M.; Maloney, D. J.; Shinn, P.; Leister, W.; Marugan, J. J.; Inglese, J.; Austin, C. P.; Misteli, T.; Auld, D. S.; Thomas, C. J. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6700; (g) Okamoto, M.; Takayama, K.; Shimizu, T.; Ishida, K.; Takahashi, O.; Furuya, T. *J. Med. Chem.* **2009**, *52*, 7323; (h) Cozza, G.; Bonvini, P.; Zorzi, E.; Poletto, G.; Pagano, M. A.; Sarno, S.; Donella-Deana, A.; Zagotto, G.; Rosolen, A.; Pinna, L. A.; Meggio, F.; Moro, S. *J. Med. Chem.* **2006**, *49*, 2363; (i) Boulahjar, R.; Ouach, A.; Matteo, C.; Bour, S.; Ravache, M.; Guevel, R. L.; Marionneau, S.; Oullier, T.; Lozach, O.; Meijer, L.; Gugen-Guillouzo, C.; Lazar, S.; Aksira, M.; Troin, Y.; Guillaumet, G.; Routier, S. *J. Med. Chem.* **2012**, *55*, 9589; (j) Frederick, R.; Bruyere, C.; Vancraeynest, C.; Reniers, J.; Meinguet, C.; Pochet, L.; Backlund, A.; Masereel, B.; Kiss, R.; Wouters, J. *J. Med. Chem.* **2012**, *55*, 6489; (k) Pagano, M. A.; Andrzejewska, M.; Ruzzene, M.; Sarno, S.; Cesaro, L.; Bain, J.; Elliott, M.; Meggio, F.; Kazmierczuk, Z.; Pinna, L. A. *J. Med. Chem.* **2004**, *47*, 6239; (l) Rosenthal, A. S.; Tanega, C.; Shen, M.; Mott, B. T.; Bougie, J. M.; Nguyen, D.-T.; Misteli, T.; Auld, D. S.; Maloney, D. J.; Thomas, C. J. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3152; (m) Demange, L.; Abdellah, F. N.; Lozach, O.; Ferandin, Y.; Gresh, N.; Meijer, L.; Galons, H. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 125; (n) Koo, K. A.; Kim, N. D.; Chon, Y. S.; Jung, M.-S.; Lee, B.-J.; Kim, J. H.; Song, W.-J. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2324; (o) Cuny, G. D.; Ulyanova, N. P.; Patnaik, D.; Liu, J.-F.; Lin, X.; Auerbach, K.; Ray, S. S.; Xian, J.; Glicksman, M. A.; Stein, R. L.; Higgins, J. M. *G. Bioorg. Med. Chem. Lett.* **2015**, *2012*, 22; (p) Chioua, M.; Samadi, A.; Soriano, E.; Lozach, O.; Meijer, L.; Marco-Contelles, J. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4566.
- Accuracy = m/N , where m is the number of actives that do match the query and in-actives that do not match the query, N is the number of molecules in the training set ($N = 86$).
- Neudert, G.; Klebe, G. *J. Chem. Inf. Model.* **2011**, *51*, 2731.
- See Supplementary material for assay details.
- (a) Zhang, G.; Cao, R.; Guo, L.; Ma, Q.; Fan, W.; Chen, X.; Li, J.; Shao, G.; Qiu, L.; Ren, Z. *Eur. J. Med. Chem.* **2013**, *65*, 21; (b) Shi, B.; Cao, R.; Fan, W.; Guo, L.; Ma, Q.; Chen, X.; Zhang, G.; Qiu, L.; Song, H. *Eur. J. Med. Chem.* **2013**, *60*, 10.
- (a) Knop, K.; Hoogenboom, R.; Fischer, D.; Schubert, U. S. *Angew. Chem.* **2010**, *122*, 6430; (b) Veronese, F. M.; Mero, A. *BioDrugs* **2008**, *22*, 315.
- Ponce, M. A.; Tarzi, O. I.; Erra-Balsells, R. *J. Heterocycl. Chem.* **2003**, *40*, 419.
- (a) El Ashry, E. S. H.; Ramadan, E. S.; Hamid, H. M. A.; Hagar, M. *Synlett* **2004**, 723; (b) Mason, J. J.; Janosik, T.; Bergman, J. *Synthesis* **2009**, 2009, 3642.
- Son, S.-Y.; Ma, J.; Kondou, Y.; Yoshimura, M.; Yamashita, E.; Tsukihara, T. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5739.
- Jain, P.; Karthikeyan, C.; Moorthy, N. S.; Walker, D. K.; Jain, A. K.; Trivedi, P. *Curr. Drug Targets* **2014**, *15*, 539.
- Grabher, P.; Durieu, E.; Kouloura, E.; Halabalaki, M.; Skaltsounis, L. A.; Meijer, L.; Hamburger, M.; Potterat, O. *Planta Med.* **2012**, *78*, 951.
- Giraud, F.; Alves, G.; Debiton, E.; Nauton, L.; Théry, V.; Durieu, E.; Ferandin, Y.; Lozach, O.; Meijer, L.; Anizon, F.; Pereira, E.; Moreau, P. *J. Med. Chem.* **2011**, *54*, 4474.

Supplementary material for

Computational & experimental evaluation of the structure/activity relationship of β -carboline as DYRK1A inhibitors.

Binia Drung ^{a,†}, Christoph Scholz ^{a,†}, Valéria A. Barbosa ^b, Azadeh Nazari ^a,
Maria H. Sarragiotto ^b and Boris Schmidt ^{a,*}.

^a*Clemens Schöpf - Institute of Organic Chemistry and Biochemistry, Technische Universität
Darmstadt, Alarich-Weiss-Str. 4, 64287 Darmstadt, Germany*

^b*Departamento de Química, Universidade Estadual de Maringá, Avenida Colombo 53790,
PR 87020-900 Maringá, Brazil*

Table of Contents:

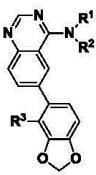
S2	I.	Pharmacophore modeling
S4	II.	Additional figures
S5	III.	Syntheses
S11	IV.	Computational procedures
S14	V.	Assay descriptions
S17	VI.	Compound characterization data
S47	VII.	References

I. Pharmacophore Modeling

Table S1. All seven elucidated five-feature pharmacophore models with their respective statistical descriptors.

	mol	label	cover	overlap	accuracy	p	acc1	acc0	query	Qsize	Q:Aro PIR	Q:Hyd	Q:Acc2
1	RRHHa_3	R R H H a	15	12.8323	0.8140	-5.1732	0.8333	0.8088	5 1 1 2 2	5	2	2	1
2	RRHHa_5	R R H H a	14	12.7675	0.8140	-4.6855	0.7778	0.8235	5 1 1 2 2	5	2	2	1
3	RRHHa_4	R R H H a	15	13.0693	0.8023	-4.8261	0.8333	0.7941	5 1 1 2 2	5	2	2	1
4	RRHHa_2	R R H H a	14	12.6468	0.8023	-4.3392	0.7778	0.8088	5 1 1 2 2	5	2	2	1
5	RRHHa_7	R R H H a	14	12.6852	0.7907	-4.0191	0.7778	0.7941	5 1 1 2 2	5	2	2	1
6	RRHHa_1	R R H H a	15	13.1592	0.7558	-3.6584	0.8333	0.7353	5 1 1 2 2	5	2	2	1
7	RRHHa_6	R R H H a	14	12.6829	0.7558	-3.1913	0.7778	0.7500	5 1 1 2 2	5	2	2	1

Table S2. Structures of active test and training set compounds and their biologic activity against DYRK1A.

	R ¹	R ²	R ³	ID	IC ₅₀ [nM]	Set
	Me		H	1 ²	14	training
	H		H	2 ²	17	test
	Me		H	3 ²	26	training
	H		H	4 ²	30	training
	Me		H	5 ²	31	test
	H		H	6 ²	35	training
	Et		H	7 ²	38	training
	Me		H	8 ²	51	test
	H		H	9 ²	57	training
	H		H	10 ²	70	test
	H		H	11 ²	74	training
	H		H	12 ²	76	training
	H		H	13 ²	84	test
	H		H	14 ²	93	training
	H		H	15 ²	93	training
	Me		H	16 ²	93	test
	Me		H	17 ²	120	test
	H		H	18 ²	126	training
	H		H	19 ²	135	training
	H		H	20 ²	155	test
	H		F	21 ²	164	training
	H		H	22 ²	180	training
	H		H	23 ²	206	test
				24 ¹	12	training
				25 ³	60	training
				26 ⁴	110	training
				27 ⁵	120	training

II. Additional figures



Figure S1. Estimated binding mode of N-heptylharminine **16**/DYRK1A obtained by docking.

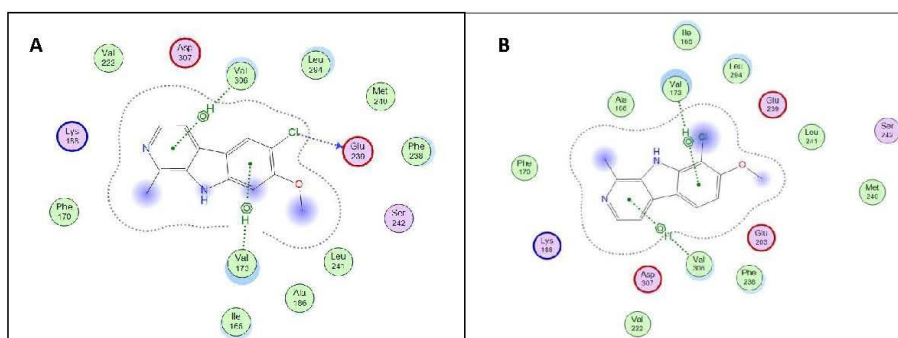


Figure S2. Protein/ligand docking of (A) 6-chloro-harminine **17** with additional interaction with Glu239 in comparison to (B) the 8-chloro-harminine **18**.

III. Syntheses

General methods: Silica gel chromatography was carried out using Merck silica gel 60 (0.015 - 0.040 mm). Thin layer chromatography was carried out using Merck pre-coated aluminium sheets silica gel 60 F₂₅₄. The chromatographic spots were indicated by fluorescence detection (UV-light, $\lambda = 254$ nm) and/or diverse staining solutions (vanillin, ninhydrin, potassium permanganate) with visualization by heating.⁶ Automated flash chromatography was performed on a Teledyne ISCO Combi Flash RF 4X using water and acetonitrile solvent mixtures. The ¹H-NMR spectra were recorded on a Bruker AC 300 spectrometer at 300 MHz and Bruker AC 500 spectrometer at 500 MHz. The ¹³C-NMR spectra were recorded on a Bruker AC 300 spectrometer at 75 MHz and Bruker AC 500 spectrometer at 125/126 MHz. Chemical shifts are reported in ppm and are calibrated to the particular solvent residual signal. ¹H-NMR data recorded in CD₃OD or CDCl₃ are listed as residual internal CD₂HOD ($\delta = 3.31$ ppm) or CHCl₃ ($\delta = 7.26$ ppm). ¹³C-NMR data recorded in CD₃OD or CDCl₃ are calibrated to CD₃OD/CD₂HOD ($\delta = 49.0$ ppm) or CDCl₃/CHCl₃ ($\delta = 77.16$ ppm).⁷ Mass spectrometry was recorded on a Bruker-Franzen Esquire LC mass spectrometer and a MAT 95 double focussing sector field MS. High performance liquid chromatography was performed using an Agilent 1100 (column: reversed phase, Zorbax Eclipse XDB-C8, 4.6 x 150 mm; 254 nm). The mobile phase was composed of: H₂O (1% TFA) (A) and acetonitrile (B) with a gradient: 30 to 90% of acetonitrile within 12 min. All commercial available reagents and solvents were purchased at VWR, Sigma Aldrich, Alfa Aesar, Acros, and ABCR and were used as supplied. 2-(2-Methoxyethoxy)ethyl tosylate, 2-(2-(2-methoxyethoxy)ethoxy)ethyl tosylate,⁸ 6-chloro-harmine **17**, 8-chloro-harmine **18**⁹ and 6-methoxyisatin **19**¹⁰ were prepared according to literature procedure. The biological assay of DYRK1A inhibition was carried out by Cerep using the a published procedure with staurosporine **3** as reference compound.¹¹ The biological assay of MAO-A inhibition was performed by Cerep using elorgyline **19** as control.¹²

General procedure A for the synthesis of β -carbolines (6-10): Tryptophane or 6-methoxytryptophane (1.0 equiv) was dissolved in acetic acid (14 mL/mmol) and the corresponding aldehyde (1.2 equiv) was added. The mixture was stirred at 70 °C for 2-3 h. The precipitate was filtered and washed with water. After concentration *in vacuo* the intermediate 1,2,3,4-tetrahydro- β -carboline-3-carboxylic acids (**5a-e**) were obtained as solids.¹³ The compounds were used for the next step without structure determination. The corresponding 1,2,3,4-tetrahydro- β -carboline-carboxylic acid (1 equiv) was suspended in an aqueous potassium dichromate solution (10% w/v, 8 mL/mmol) and glacial acetic acid. The mixture was refluxed for 3-6 h. After cooling to rt Na₂SO₃ (aq, saturated) was added for reduction of the oxidant excess and the mixture was neutralized with Na₂CO₃. The aqueous layer was extracted three times with DCM. The combined organic layers were dried over MgSO₄ and concentrated *in vacuo*.

7-Methoxy-1-ethyl- β -carboline 6: Yield 37% after purification by silica gel chromatography (ethyl acetate), yellow solid. ¹H-NMR (CD₃OD, 300 MHz) δ = 8.16 (d, J = 5.3 Hz, 1H), 8.05 (d, J = 8.7 Hz, 1H), 7.93 (d, J = 5.1 Hz, 1H), 7.08 (d, J = 2.0 Hz, 1H), 6.92 (dd, J = 8.8, 2.0 Hz, 1H), 3.94 (s, 3H), 3.20 (q, J = 7.5 Hz, 2H), 1.45 (t, J = 7.5 Hz, 3H) ppm. ¹³C-NMR (CD₃OD, 75 MHz) δ = 163.3, 146.3, 145.1, 136.2, 135.3, 131.7, 123.9, 116.2, 113.8, 111.7, 95.4, 56.1, 27.0, 13.6 ppm. MS (m/z , 70 eV, EI) = 226 [M]⁺.

7-Methoxy-1-propyl- β -carboline 7: Yield 40% after purification by silica gel chromatography (ethyl acetate), yellow solid. ¹H-NMR (CD₃OD, 300 MHz) δ = 8.12 (d, J = 5.5 Hz, 1H), 7.99 (d, J = 8.7 Hz, 1H), 7.82 (d, J = 5.5 Hz, 1H), 7.04 (d, J = 2.2 Hz, 1H), 6.87 (dd, J = 8.7, 2.3 Hz, 1H), 3.90 (s, 3H), 3.15-2.99 (m, 2H), 1.98 – 1.75 (m, 2H), 1.03 (t, J = 7.4 Hz, 3H) ppm. ¹³C-NMR (CD₃OD, 75 MHz) δ = 164.3, 147.5, 145.9, 139.2, 132.3, 125.1, 117.8, 114.9, 112.5, 101.4, 96.9, 57.5, 37.8, 25.0, 15.7 ppm. MS (m/z , 70 eV, EI) = 240 [M]⁺.

Harmane 8: Yield: 53% after recrystallization from ethyl acetate, yellow solid. $^1\text{H-NMR}$ (CD_3OD , 300 MHz) δ = 8.15 - 8.11 (m, 2H), 7.91 (d, J = 5.4 Hz, 1H), 7.60 – 7.51 (m, 2H), 7.24 (ddd, J = 8.1, 6.4, 1.7 Hz, 1H), 2.80 (s, 3H) ppm. $^{13}\text{C-NMR}$ (CD_3OD , 75 MHz) δ = 142.8, 137.2, 136.2, 130.1, 129.6, 122.7, 122.6, 120.9, 114.2, 112.9, 19.4 ppm. **MS** (m/z , 70 eV, EI) = 182 $[\text{M}]^+$.

1-Ethyl- β -carboline 9: Yield 51% after recrystallization from ethyl acetate, yellow solid. $^1\text{H-NMR}$ (CD_3OD , 300 MHz) δ = 8.15 (d, J = 5.5 Hz, 1H), 8.13 – 8.07 (dt, J = 7.9, 1.0 Hz, 1H), 7.87 (d, J = 5.4 Hz, 1H), 7.59 – 7.48 (m, 2H), 7.21 (ddd, J = 8.0, 6.7, 1.4 Hz, 1H), 3.14 (q, J = 7.6 Hz, 2H), 1.40 (t, J = 7.6 Hz, 3H) ppm. $^{13}\text{C-NMR}$ (CD_3OD , 75 MHz) δ = 148.3, 142.6, 137.8, 135.4, 130.2, 129.4, 122.6, 122.6, 120.7, 114.1, 112.8, 27.8, 13.6 ppm. **MS** (m/z , 70 eV, EI) = 196 $[\text{M}]^+$.

1-Isobutyl- β -carboline 10: Yield 34% after recrystallization from ethyl acetate, yellow solid. $^1\text{H-NMR}$ (CDCl_3 , 300 MHz) δ = 9.11 (bs, 1H), 8.32 (d, J = 5.4 Hz, 1H), 8.04 (d, J = 7.9 Hz, 1H), 7.75 (d, J = 5.4 Hz, 1H), 7.48 – 7.37 (m, 2H), 7.25 – 7.14 (m, 1H), 2.93 (d, J = 7.4 Hz, 2H), 2.36 – 2.17 (m, 1H), 0.88 (d, J = 6.6 Hz, 6H) ppm. $^{13}\text{C-NMR}$ (CDCl_3 , 75 MHz) δ = 145.2, 140.4, 138.2, 134.8, 128.8, 128.3, 121.9, 121.8, 120.0, 112.9, 111.6, 43.2, 28.8, 22.8 ppm. **MS** (m/z , 70 eV, EI) = 224 $[\text{M}]^+$.

General procedure B for synthesis of 7-substituted-1-methyl-9H-pyrido[3,4-b]indoles (12-15): Harmol **2** (1 equiv) and cesium carbonate (1.5 equiv) were suspended in anhydrous DMF (20 mL/mmol) under argon atmosphere. The mixture was pre-activated for 30 min at rt. After cooling to 0 °C the appropriate alkylation/pegylation reagent (1.1 equiv - 1.3 equiv) was added and the mixture was stirred at 0 °C for 1 h. The suspension was warmed up to rt and stirred over night. After addition of water, the aqueous layer was extracted three times with ethyl acetate. The combined organic layers were dried over MgSO_4 and concentrated

in vacuo. Purification was done by flash chromatography (RediSep Column: 13 g C18 gold, water/acetonitrile = 0% → 95% acetonitrile in 12 min).

7-Ethoxy-1-methyl-9H-pyrido[3,4-b]indole 12:¹⁴ Yield 72%, colorless solid. ¹H-NMR (CDCl₃, 500 MHz): δ = 8.43 (bs, 1H), 8.32 (d, *J* = 5.4 Hz, 1H), 7.95 (d, *J* = 8.6 Hz, 1H), 7.71 (d, *J* = 5.3 Hz, 1H), 6.94 (d, *J* = 2.1 Hz, 1H), 6.90 (dd, *J* = 8.6, 2.2 Hz, 1H), 4.12 (q, *J* = 7.0 Hz, 2H), 2.79 (s, 3H), 1.47 (t, *J* = 7.0 Hz, 3H) ppm. ¹³C-NMR (CDCl₃, 126 MHz): δ = 160.4, 141.9, 141.1, 139.1, 134.8, 128.9, 122.8, 116.0, 112.3, 110.3, 95.6, 64.0, 20.3, 15.0 ppm. MS (*m/z*, 70 eV, EI) = 226 [M]⁺.

7-Heptyloxy-1-methyl-9H-pyrido[3,4-b]indole 13: Yield 50%, colorless solid. ¹H-NMR (CDCl₃, 500 MHz): δ = 9.37 (bs, 1H), 8.26 (d, *J* = 5.5 Hz, 1H), 7.95 (d, *J* = 8.7 Hz, 1H), 7.74 (d, *J* = 5.5 Hz, 1H), 7.01 (d, *J* = 2.1 Hz, 1H), 6.91 (dd, *J* = 8.7, 2.2 Hz, 1H), 4.03 (t, *J* = 6.6 Hz, 2H), 2.87 (s, 3H), 1.89-1.76 (m, 2H), 1.54-1.42 (m, 2H), 1.42-1.22 (m, 6H), 0.91 (t, *J* = 7.0 Hz, 3H) ppm. ¹³C-NMR (CDCl₃, 126 MHz): δ = 161.1, 142.8, 140.2, 136.8, 134.8, 129.7, 122.8, 115.5, 112.5, 111.1, 95.6, 68.7, 31.9, 29.4, 29.2, 26.2, 22.8, 19.5, 14.2 ppm. MS (*m/z*, 70 eV, EI) = 296 [M]⁺.

7-(2-(2-Methoxyethoxy)ethoxy)-1-methyl-9H-pyrido[3,4-b]indole 14: Yield 34%, colorless solid. ¹H-NMR (CD₃OD, 500 MHz): δ = 8.08 (d, *J* = 5.4 Hz, 1H), 7.95 (d, *J* = 8.6 Hz, 1H), 7.76 (dd, *J* = 5.5, 0.7 Hz, 1H), 7.03 (d, *J* = 2.1 Hz, 1H), 6.86 (dd, *J* = 8.7, 2.2 Hz, 1H), 4.23 - 4.18 (m, 2H), 3.90 - 3.85 (m, 2H), 3.74 - 3.68 (m, 2H), 3.60 - 3.54 (m, 2H), 3.36 (s, 3H), 2.74 (s, 3H) ppm. ¹³C-NMR (CD₃OD, 125 MHz): δ = 161.6, 144.1, 142.1, 138.0, 136.3, 130.1, 123.5, 116.6, 113.3, 111.2, 96.4, 73.0, 71.6, 70.8, 68.9, 59.1, 19.5 ppm. MS (*m/z*, 70 eV, EI) = 300 [M]⁺.

7-(2-(2-(2-Methoxyethoxy)ethoxy)ethoxy)-1-methyl-9H-pyrido[3,4-b]indole 15:

Yield 41%, colorless solid. ¹H-NMR (CD₃OD, 500 MHz): δ = 8.08 (d, *J* = 5.3 Hz, 1H), 7.96 (d, *J* = 8.7 Hz, 1H), 7.76 (t, *J* = 3.8 Hz, 1H), 7.04 (d, *J* = 2.3 Hz, 1H), 6.89 - 6.83 (m, 1H),

4.24 – 4.22 (m, 2H), 3.91 - 3.84 (m, 2H), 3.74 - 3.57 (m, 6H), 3.52 - 3.46 (m, 2H), 3.33 (s, 3H), 2.74 (d, $J = 3.1$ Hz, 3H) ppm. $^{13}\text{C-NMR}$ (CD_3OD , 125 MHz): $\delta = 161.7, 144.1, 142.1, 138.0, 136.3, 130.1, 123.6, 116.6, 113.3, 111.3, 96.5, 73.0, 71.8, 71.6, 71.4, 70.9, 69.0, 59.1, 19.5$ ppm. **MS** (m/z , 70 eV, EI) = 344 $[\text{M}]^+$.

9-Heptyl-harmine 16: NaH (10 mg, 0.19 mmol, 55% dispersion) was suspended in anhydrous DMF under argon atmosphere and harmine **1** (20 mg, 0.09 mmol) was added. After 30 min pre-activation the mixture was cooled to 0 °C and 1-iodoheptane (20.1 μL , 0.12 mmol) was added. The reaction mixture was stirred at 0 °C for 1 h. The suspension was warmed up to rt and stirred over night. After addition of water the aqueous layer was extracted three times with ethyl acetate. The combined organic layers were dried over MgSO_4 and concentrated *in vacuo*. After purification by flash chromatography (RediSep Column: 13 g C18 gold, water : acetonitrile = 0% \rightarrow 95% acetonitrile in 12 min) the product **16** was

obtained in 95% yield as a colorless solid. $^1\text{H-NMR}$ (CDCl_3 , 500 MHz): $\delta = 8.28$ (d, $J = 5.4$ Hz, 1H), 7.99 (d, $J = 8.6$ Hz, 1H), 7.78 (d, $J = 8.6$ Hz, 1H), 6.91 (dd, $J = 8.6, 2.2$ Hz, 1H), 6.86 (d, $J = 2.1$ Hz, 1H), 4.48 - 4.44 (m, 2H), 3.96 (s, 2H), 3.07 (s, 3H), 1.83 (dt, $J = 15.5, 7.7$ Hz, 2H), 1.45-1.27 (m, 8H), 0.88 (t, $J = 6.9$ Hz, 3H) ppm. $^{13}\text{C-NMR}$ (CDCl_3 , 126 MHz): $\delta = 161.4, 143.7, 140.1, 137.0, 135.3, 130.2, 122.8, 115.2, 112.6, 109.3, 93.6, 55.9, 45.2, 31.8, 30.8, 29.0, 27.0, 22.7, 14.2$ ppm. **MS** (m/z , 70 eV, EI) = 310 $[\text{M}]^+$.

5-Isopentyl-7-methoxy-3-(methylthio)-5H-[1,2,4]triazino[5,6-b]indole 22:

6-Methoxyisatin **19** (150 mg, 0.85 mmol) was dissolved in DMF (10 mL), K_2CO_3 (234 mg, 1.70 mmol) and 3-methylbutyl bromide (211 μL , 1.70 mmol) were added. The reaction mixture was stirred at rt for 24 h. After addition of water the precipitate was filtered off and dried *in vacuo*. The intermediate 1-isopentyl-6-methoxyindoline-2,3-dione **20a** was obtained in 74% yield. **20a** (100 mg, 0.40 mmol) and thiosemicarbazide (52 mg, 0.66 mmol) were suspended in ethanol (10 mL) with a drop of acetic acid. The mixture was irradiated by

microwave at 80 °C for 2 h. The solvent was removed *in vacuo* and the residue was suspended with K₂CO₃ (67 mg, 0.48 mmol) in water (10 mL). The mixture was re-irradiated at 100 °C for 2.5 h. The reaction mixture was acidified with acetic acid (pH = 2) and filtrated. The filtrate was washed with water and dried *in vacuo* to obtain the intermediate 5-isopentyl-7-methoxy-5*H*-[1,2,4]triazino[5,6-*b*]indole-3-thiol **21a** in 50% yield. To a solution of **21a** (50 mg, 0.16 mmol) in DMF (5 mL), K₂CO₃ (46 mg, 0.31 mmol) and iodomethane (21 µL, 0.31 mmol) were added. The reaction mixture was stirred at rt for 8 h. After addition of water the precipitate was filtered off and dried *in vacuo*. The compound **22** was obtained as yellow solid in 69% yield. ¹H-NMR (CDCl₃, 500 MHz): δ= 8.31 (d, *J* = 8.5 Hz, 1H), 7.02 (dd, *J* = 8.5, 2.2 Hz, 1H), 6.92 (d, *J* = 1.9 Hz, 1H), 4.32 (t, *J* = 7.4 Hz, 2H), 3.99 (s, 3H), 2.77 (s, 3H), 1.77 (q, *J* = 6.7 Hz, 2H), 1.59 – 1.69 (m, 1H), 1.04 (d, *J* = 6.6 Hz, 6H) ppm. ¹³C-NMR (CDCl₃, 125 MHz): δ= 167.1, 162.6, 147.1, 142.9, 140.7, 123.7, 111.4, 109.9, 95.6, 55.8, 39.7, 36.8, 25.8, 22.4, 14.1 ppm. MS (*m/z*, 70 eV, EI) = 316 [M]⁺.

5-Isopentyl-8-methoxy-3-(methylthio)-5*H*-[1,2,4]triazino[5,6-*b*]indole 23:

5-Methoxyisatin **24** (195 mg, 1.1 mmol) was dissolved in DMF (10 mL), K₂CO₃ (30 mg, 2.2 mmol) and 3-methylbutyl bromide (277 µL, 2.2 mmol) were added. The reaction mixture was stirred at rt for 24 h. After addition of water the precipitate was filtered off and dried *in vacuo*. The intermediate 1-isopentyl-5-methoxyindoline-2,3-dione **20b** was obtained in 95% yield. **20b** (247 mg, 1.0 mmol) and thiosemicarbazide (103 mg, 1.1 mmol) were suspended in ethanol (10 mL) with a drop of acetic acid. The mixture was irradiated by microwave at 80 °C for 2 h. The solvent was removed *in vacuo* and the residue was suspended with K₂CO₃ (138 mg, 1.0 mmol) in water (10 mL). The mixture was re-irradiated at 100 °C for 2.5 h. The reaction mixture was acidified with acetic acid (pH = 2) and filtrated. The filtrate was washed with water and dried *in vacuo* to obtain the intermediate 5-isopentyl-8-methoxy-5*H*-[1,2,4]triazino[5,6-*b*]indole-3-thiol **21b** in 71% yield. To a solution of **21b** (200 mg, 0.7 mmol) in DMF (5 mL), K₂CO₃ (180 mg, 1.3 mmol) and iodomethane (81 µL, 1.3 mmol)

were added. The reaction mixture was stirred at rt for 8 h. After addition of water the precipitate was filtered off and dried *in vacuo*. The compound **23** was obtained as a yellow solid in 63% yield. ¹H-NMR (CDCl₃, 500 MHz): δ = 7.93 (d, J = 2.2 Hz, 1H), 7.39 (d, J = 8.8 Hz, 1H), 7.29 - 7.32 (m, 1H), 4.36 (t, J = 7.4 Hz, 2H), 3.97 (s, 3H), 2.77 (s, 3H), 1.78 (q, J = 6.9 Hz, 2H), 1.62 (n, J = 6.3 Hz, 1H), 1.04 (d, J = 6.6 Hz, 6H) ppm. ¹³C-NMR (CDCl₃, 125 MHz): δ = 168.2, 156.2, 146.4, 140.6, 135.3, 119.8, 119.0, 111.1, 104.9, 56.0, 39.8, 37.0, 25.8, 22.3, 14.0 ppm. MS (m/z , 70 eV, ED) = 316 [M]⁺.

IV. Computational procedures:

Pharmacophore modeling: The pharmacophore models were built on the basis of a dataset provided by the public and web-accessible database bindingdb.org.¹⁵ The search for DYRK1A-inhibitors resulted in 228 entries with published IC₅₀-values (access date: 10/24/2013). This dataset was subdivided into 27 active (IC₅₀ ≤ 250 nM) and 100 inactive non-duplicate compounds (IC₅₀ ≥ 10 μ M), which are further annotated as actives and inactives, respectively. The active training set (two-thirds of all active compounds) and test set compounds (one-third of all active compounds) were defined in a way that they covered a similar range of activities and are shown in **Table S2**.

The conformational space of each molecule was explored with the systematic, forcefield- and fragment-based *ConfImport* application within the software MOE.¹⁶ The input molecule is therefore divided into overlapping fragments whose conformations are generated independently prior to reassembly. A maximum of 250 conformations per molecule was generated, using the *MMFF94x* forcefield with a distance-dependent dielectric for solvent treatment.¹⁷ The pharmacophore elucidator application of MOE generates possible pharmacophore queries that are shared by all active compounds. Several possible features were allowed for the elucidation, including π -rings, H-bond acceptor or donor projections and a maximum of three hydrophobic centroids. As the space of all possible queries is continuous

and impossible to explore completely, a tree-like algorithm is used, which generates larger queries from smaller ones. Only five-feature queries that cover at least 80% of the actives from the training set were generated and scored for their overall accuracy. ($\text{Acc} = m/N$, where m is the number of actives that do match the query and inactives that do not match the query, N is the number of molecules in the training set ($N = 86$)). Queries with an overall accuracy ≥ 0.8 were evaluated against test set compounds.

Preparation of the receptor and ligand molecules for docking: Ligands were sketched with the Builder tool of MOE.¹⁶ Hydrogens were added and charges assigned using the potential setup interface. Low-energy conformations of ligands were prepared via energy minimization with the *MMFF94x* forcefield and a distance-dependent dielectric for solvation treatment. The X-ray structure of the **1**/DYRK1A-complex (PDB-ID: 1ANR) and the **1**/MAO-A-complex (PDB-ID: 2Z5X) were prepared with the *LigX* interface. *LigX* combines several structure preparation steps including deletion of incidental waters, *Protonate 3D*, assigning ASN/GLN/HIS-“Flips”, geometry optimization and energy minimization (RMSD gradient = 0.1 kcal/mol) of the prepared receptor. The *Protonate 3D* application assigns protonation states from a discrete collection of states by optimizing the titration free energy of all titratable groups in the context of an all-atom model of a macromolecular structure (including ligands and solvent).¹⁸ The Generalized Born/Volume Integral Electrostatics model is used for longer range interactions and solvation effects.¹⁹ For energy minimization the *AMBER12:EHT* forcefield was used, which is parameterized for both, proteins and small molecules.²⁰

Docking: The Dock workflow in MOE is subdivided into the following steps: (1) conformation generation, (2) placement, (3) scoring, (4) pose refinement and (5) rescoring. Conformations are generated by applying a collection of preferred torsion angles to rotatable bonds. A systematic search is carried out, covering all combinations of angles on a grid if this will result in less than 5000 conformers. Otherwise a stochastic sampling of

S12

conformations is conducted. A pharmacophore-based placement was applied, utilizing one aromatic and one nitrogen positioning constraint based on the 1/DYRK1A complex and one additional aromatic positioning constraint based on the 1/MAO-A complex. The generated poses were rated by the scoring function *London dG* which comprises five components: (1) loss/gain of rotational and translational entropy, (2) energy due to loss of flexibility, (3) energy due to hydrogen bonds, (4) energy due to metal ligation and (5) desolvation energy of atoms. The 30 highest scored poses were further refined by energy minimization inside the active site using the *AMBER12:EHT* forcefield to a RMSD gradient of 0.01 kcal/mol. The poses were rescored in the same way as in the first run and the ten highest rated conformations were collected in the output database. According to our experience from previous docking projects, the most similar pose compared to the experimentally determined X-ray structure is not compellable the one with the highest score from *London dG* but it is most likely one of the top ten. Due to this issue a second rescoring of the output conformations was done with the knowledge-based scoring function $DSX^{CSD,21}$. In contrast to their empirical counterparts, e.g. *London dG*, knowledge-based scoring functions are not trained with binding affinities. These functions calculate the total score as a sum of statistical potentials, which are derived from a database of known protein/ligand complexes. As the score of *DSX* reflects native binding geometries, predicting a near native conformation is one of its key skills.

V. Assay descriptions

DYRK1A

The evaluation of the activity was performed by the company Cerep²² (ref. 2781) using a published procedure by Himpel et al.¹¹ The used enzyme is a human recombinant of DYRK1A. The substrates ATP and Ulight-CFFKNIVTPRTPPPSQGK-amide (100 nM) are incubated for 30 min at room temperature. The inhibition was determined by measurement of the substrate phosphorylation to phospho-Ulight-CFFKNIVTPRTPPPSQGK-amide in presence of a potential inhibitor in comparison to a control. For reference staurosporin **3** ($IC_{50}(\text{DYRK1A}) = 25 \text{ nM}$) was used. The detection was performed using the LANCE method, which is based on a FRET between labeled antibodies and the substrate. Each measurement was carried out as duplicates. The results from single point measurements are expressed as a percentage of control specific activity = $100 - ((\text{measured specific activity} / \text{control specific activity}) * 100)$ obtained in presence of 1 μM and/or 10 μM of the test compound. The IC_{50} value of compd. **16** was determined by measurement of eight concentrations (0,1 nM, 1 nM, 10 nM, 100 nM, 500 nM, 1 μM , 10 μM , 100 μM) under identical conditions.

MAO-A

The evaluation of the activity was performed by the company Cerep²² (ref. 0191) using a published procedure by Weyler et al.²³ The source of the MAO-A enzyme is human placenta. The activity is determined by photometry, therefore the increase of absorption is measured at a wavelength of 314 nm in presence of a potential inhibitor in comparison to a control. This is based on the oxidation of the substrat kynuramin (0.15 mM) to 4-hydroxyquinolin.²⁴ The incubation was carried out for 30 min at 30°C. For reference clorgyline **19** ($IC_{50}(\text{MAO-A}) = 37 \text{ nM}$) was used. Each measurement was carried out as duplicates. The results from single point measurements are expressed as a percentage of control specific

activity = 100-((measured specific activity/control specific activity)*100) obtained in presence of 1 μ M and/or 10 μ M of the test compound.

DYRK2

The evaluation of the activity was performed by the company Cerep²² (ref. 2788) using a published procedure by Taira et al.²⁵ Identical conditions compared to the DYRK1A assay were used. For reference staurosporin **3** (IC_{50} (DYRK2) = 1100 nM) was used. Each measurement was carried out as duplicates. The results from single point measurements are expressed as a percentage of control specific activity = 100-((measured specific activity/control specific activity)*100) obtained in presence of 1 μ M of the test compound.

DYRK3

The evaluation of the activity was performed by the company Cerep²² (ref. 1963) using a published procedure by Li et al.²⁶ The used enzyme is a human recombinant (insect cells) of DYRK3. The substrates ATP and biotinyl-MBP (40 nM) are incubated for 60 min at room temperature. The inhibition was determined by measurement of the substrate phosphorylation to phospho-biotinyl-MBP in presence of the test compound. For reference staurosporin **3** (IC_{50} (DYRK3) = 37 nM) was used. The detection was performed using the HTRF-method. Each measurement was carried out as duplicates. The results from single point measurements are expressed as a percentage of control specific activity = 100-((measured specific activity/control specific activity)*100) obtained in presence of 1 μ M of the test compound.

DYRK4

The evaluation of the activity was performed by the company Cerep²² (ref. 1962) using a published procedure by Li et al.²⁶ The used enzyme is a human recombinant (insect cells) of DYRK4. The substrates ATP and biotinyl-MBP (160 nM) are incubated for 15 min at room temperature. The inhibition was determined by measurement of the substrate phosphorylation to phospho-biotinyl-MBP in presence of a potential inhibitor in comparison to a control. For reference hymenialdisine (IC_{50} (DYRK4) = 149 nM) was used. The detection was performed

using the HTRF-method. Each measurement was carried out as duplicates. The results from single point measurements are expressed as a percentage of control specific activity = $100 - ((\text{measured specific activity} / \text{control specific activity}) * 100)$ obtained in presence of 1 μM of the test compound.

CLK1

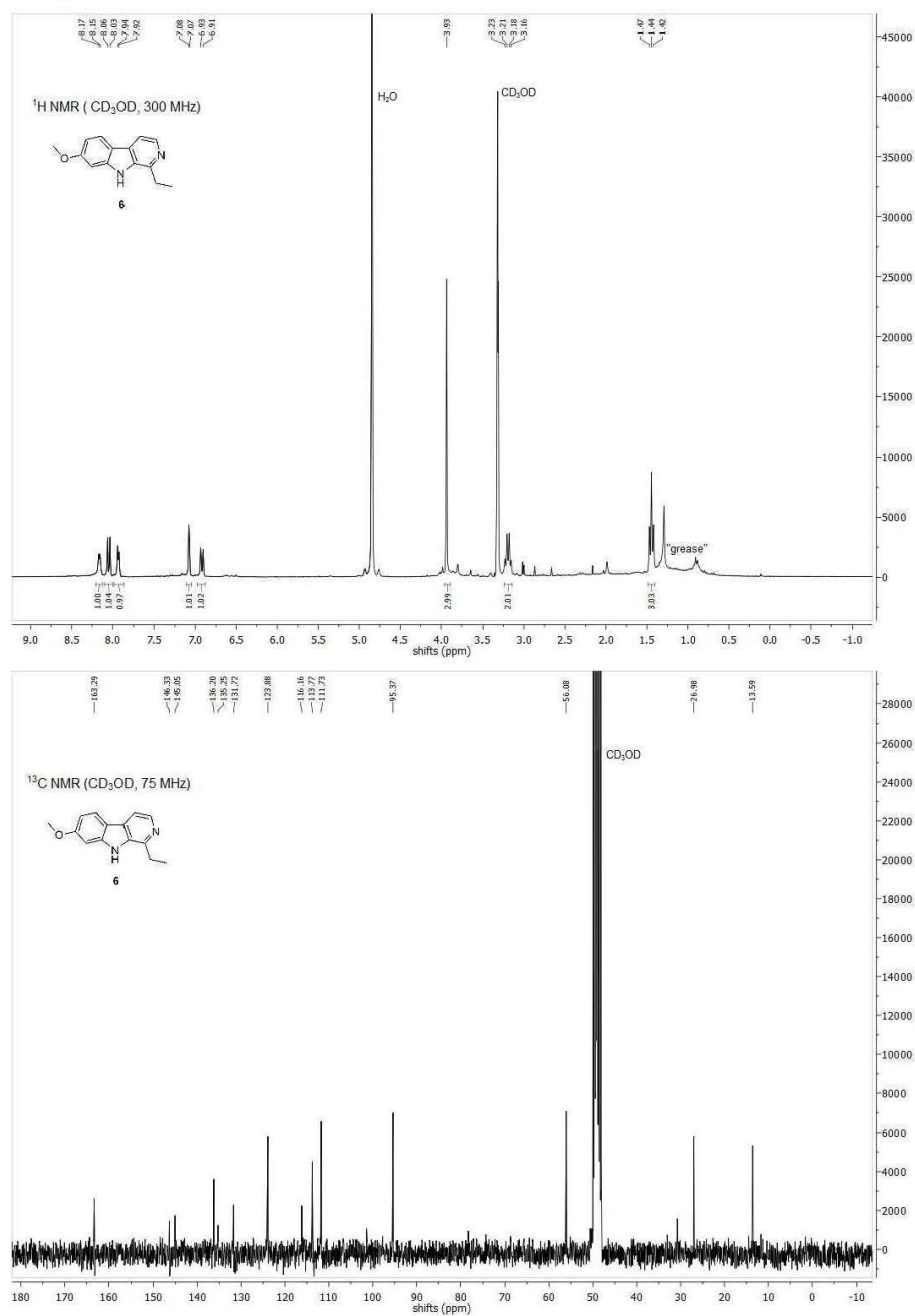
The evaluation of the activity was performed by the company Cerep²² (ref. 2787) using a published procedure by Muraki et al.²⁷ Identical conditions compared to the DYRK1A assay were used. For reference staurosporin **3** ($\text{IC}_{50}(\text{CLK1}) = 15 \text{ nM}$) was used. Each measurement was carried out as duplicates. The results from single point measurements are expressed as a percentage of control specific activity = $100 - ((\text{measured specific activity} / \text{control specific activity}) * 100)$ obtained in presence of 1 μM of the test compound.

CLK2

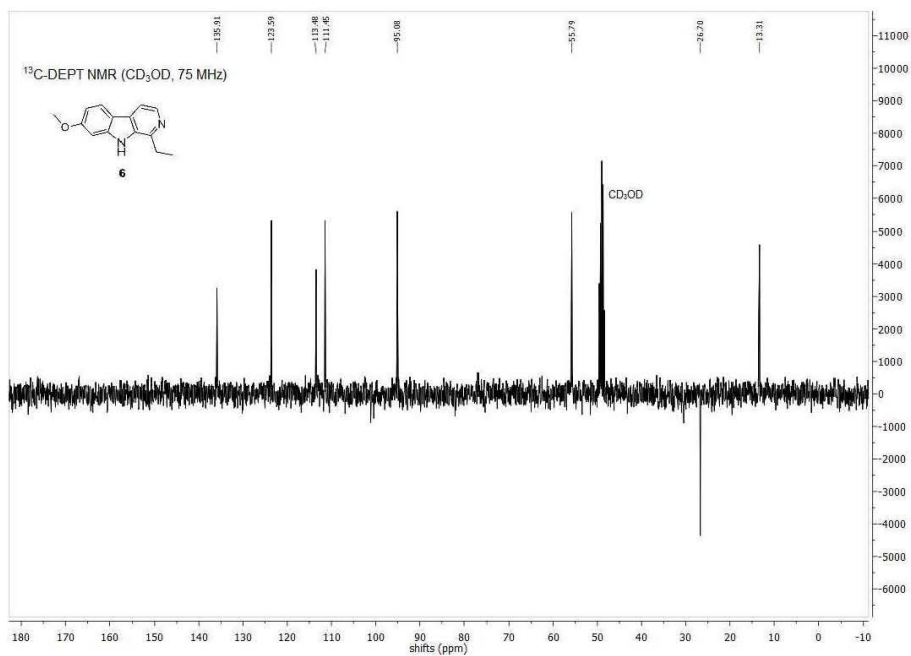
The evaluation of the activity was performed by the company Cerep²² (ref. 2019) using a published procedure by Nayler et al.²⁸ The used enzyme is a human recombinant (insect cells) of CLK2. The substrates ATP and + biotinyl-bAbAbAAGAGKRREILSRRPSYRK (10 nM) are incubated for 30 min at room temperature. The inhibition was determined by measurement of the substrate phosphorylation to phospho-biotinyl-bAbAbAAGAGKRREILSRRPSYRK in presence of the test compound in comparison to a control. For reference Ro-318220 ($\text{IC}_{50}(\text{DYRK3}) = 37 \text{ nM}$) was used. The detection was performed using the HTRF-method. Each measurement was carried out as duplicates. The results from single point measurements are expressed as a percentage of control specific activity = $100 - ((\text{measured specific activity} / \text{control specific activity}) * 100)$ obtained in presence of 1 μM of the test compound.

VI. Compound Characterization Data

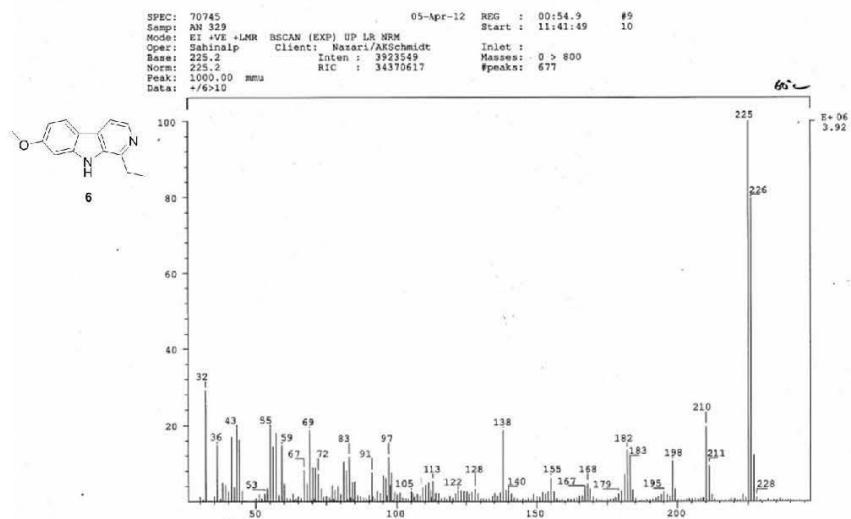
Compound 6



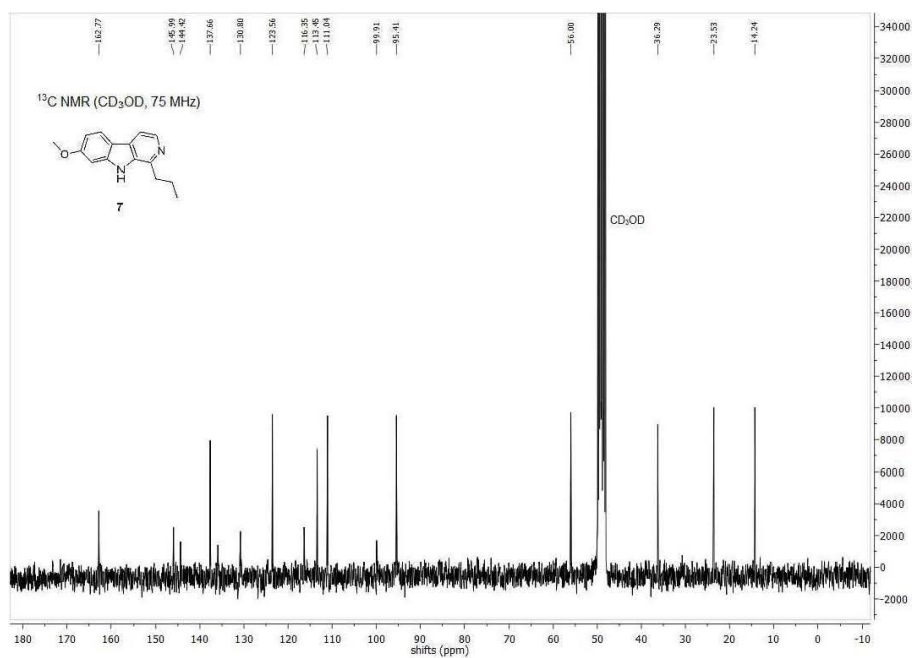
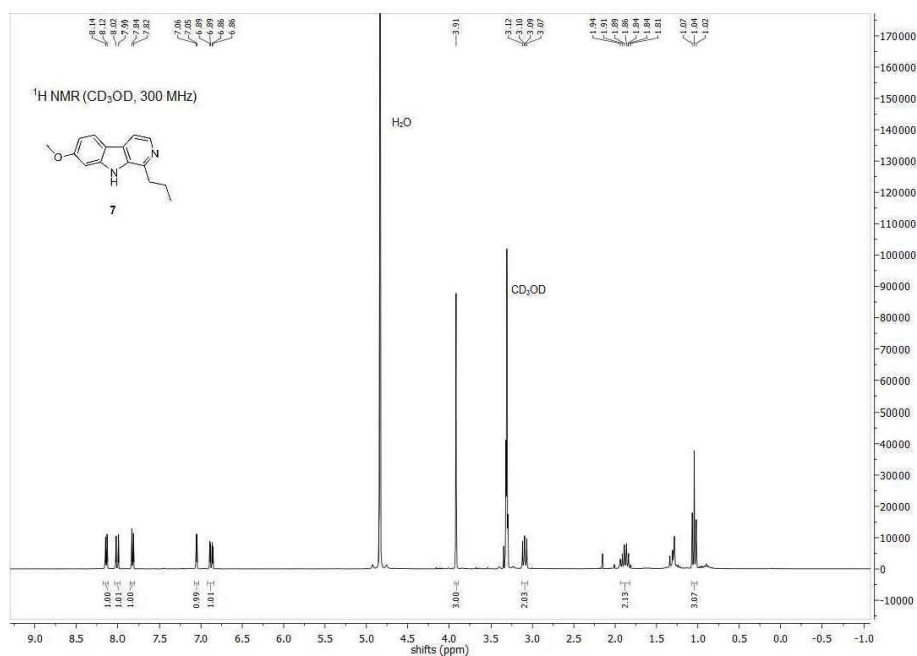
S17



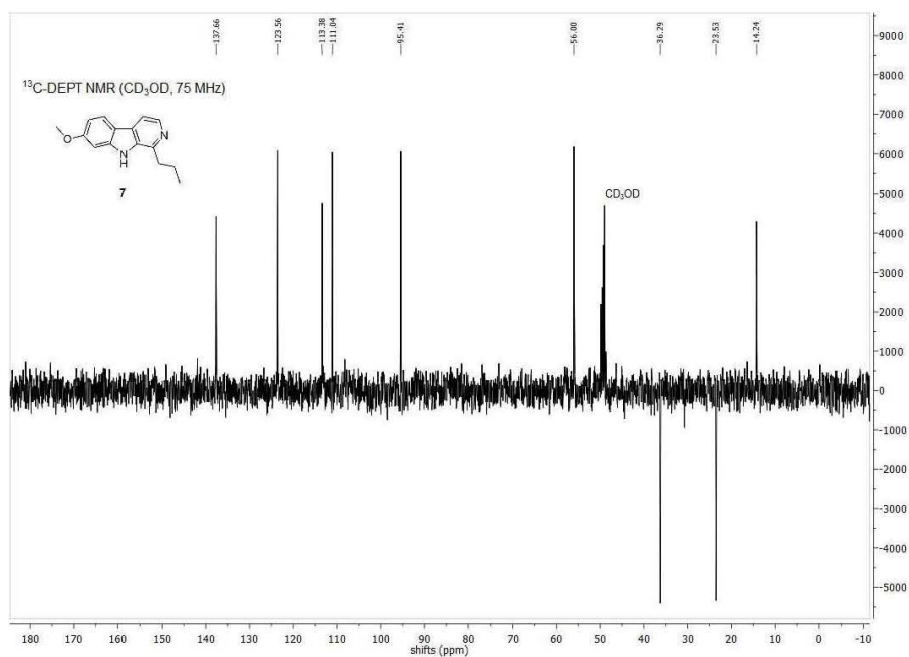
MS (m/z, 70 eV, EI)



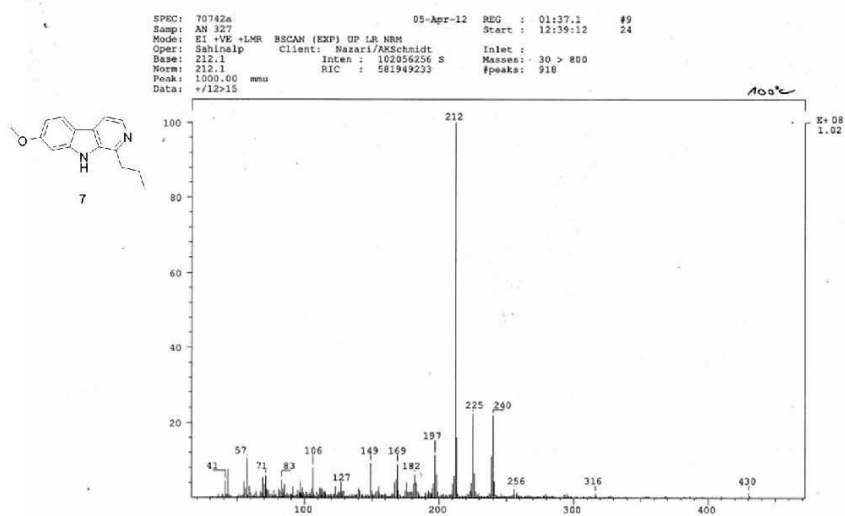
Compound 7



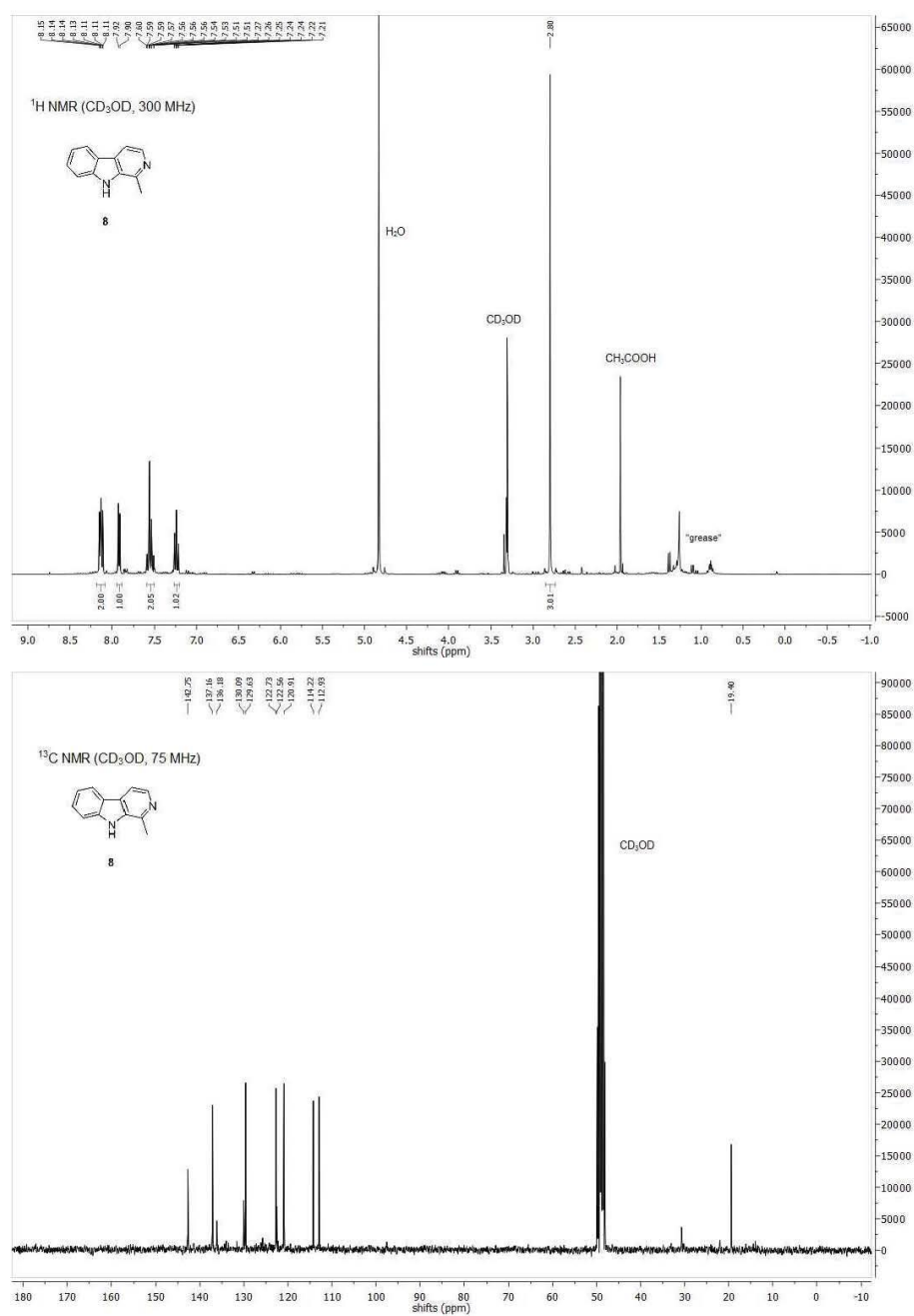
S19



MS (m/z, 70 eV, B)

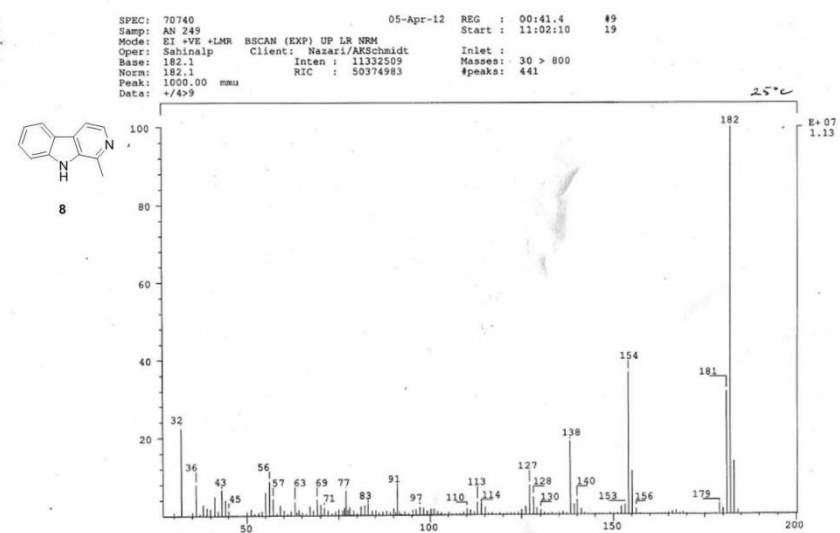


Compound 8



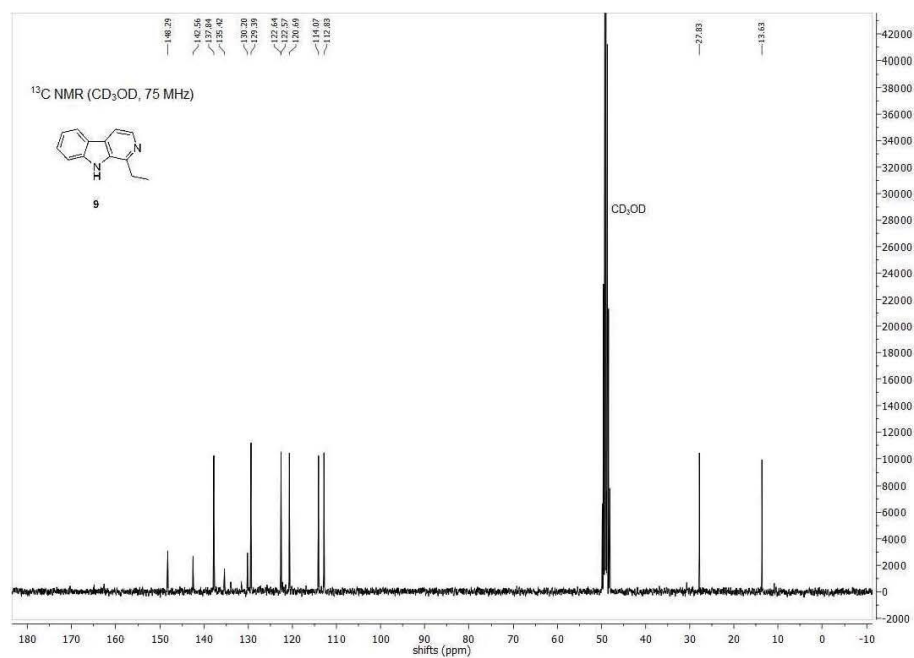
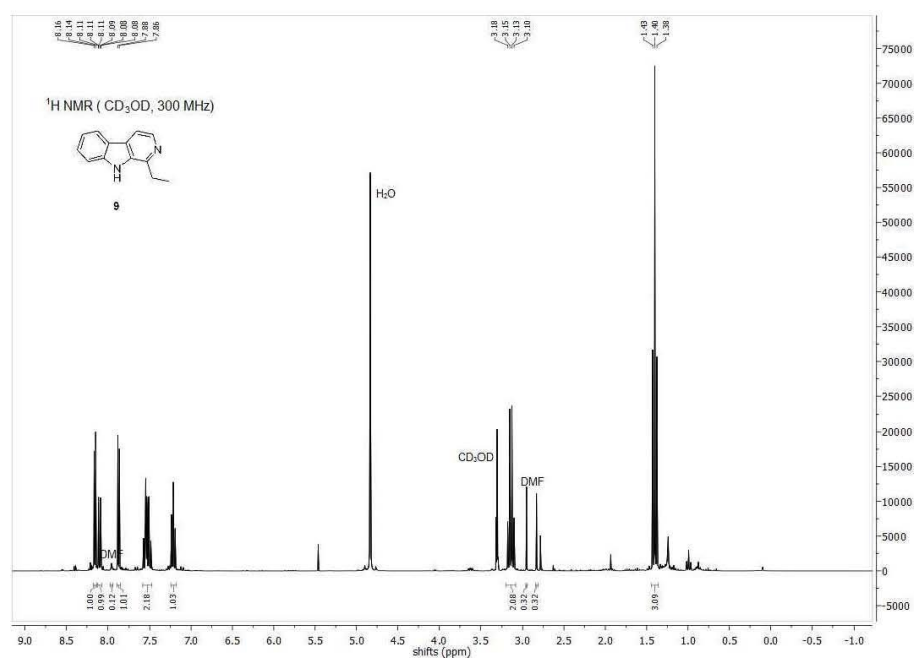
521

MS (m/z, 70 eV, EI)

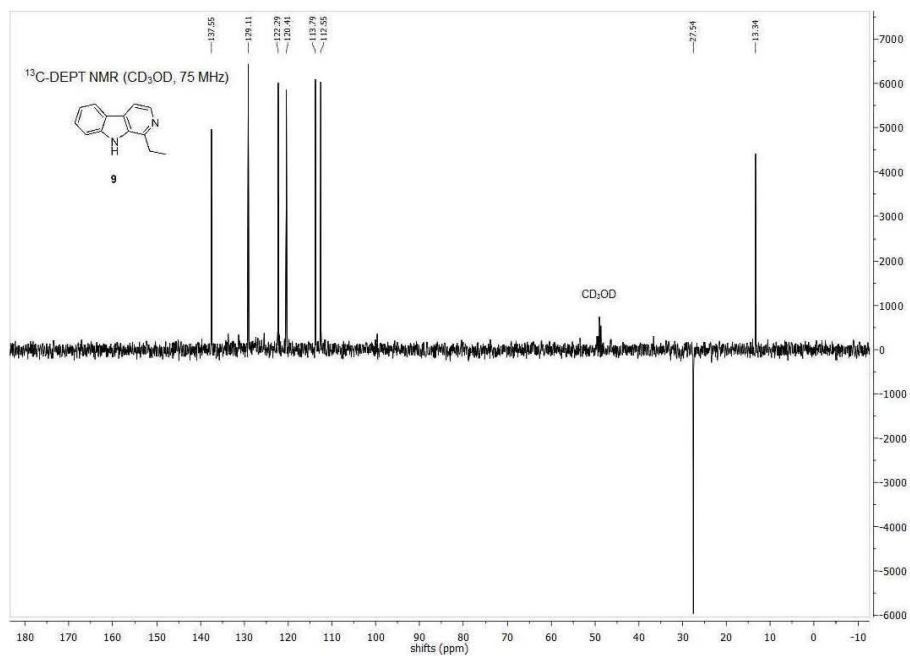


S22

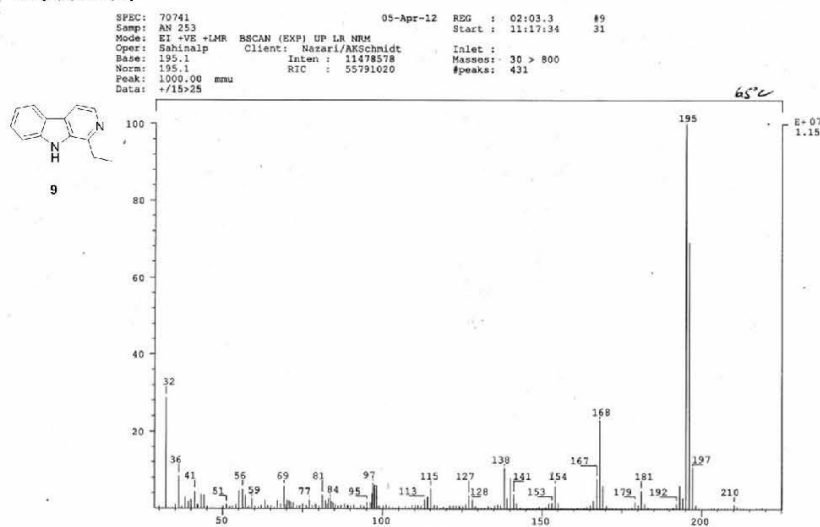
Compound 9



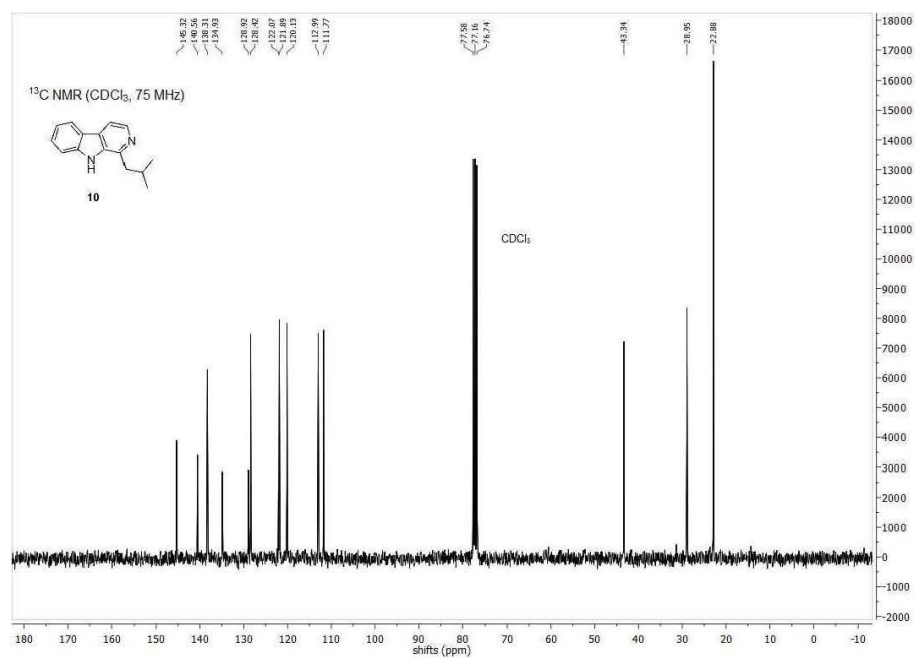
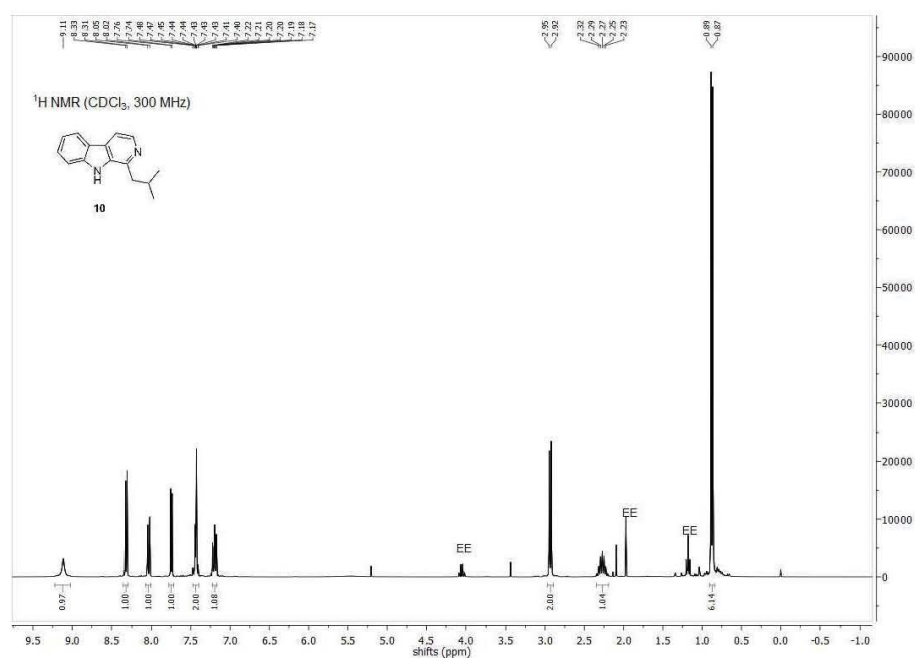
S23

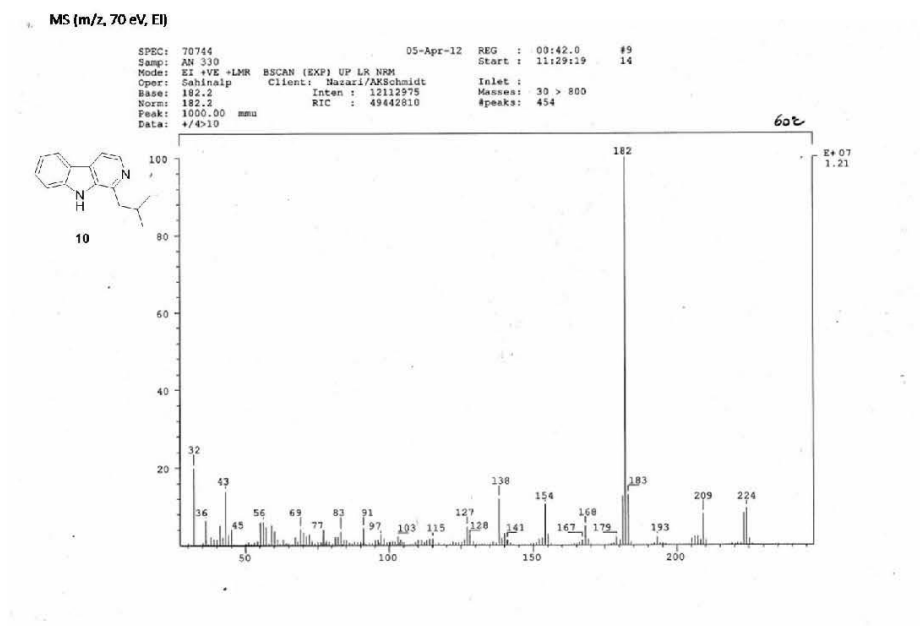
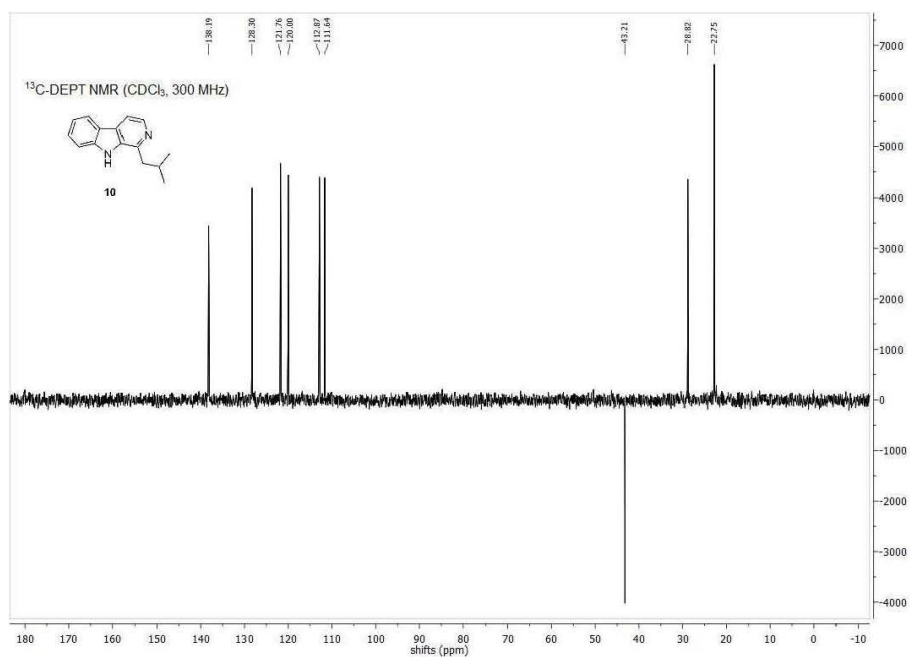


MS (m/z, 70 eV, EI)

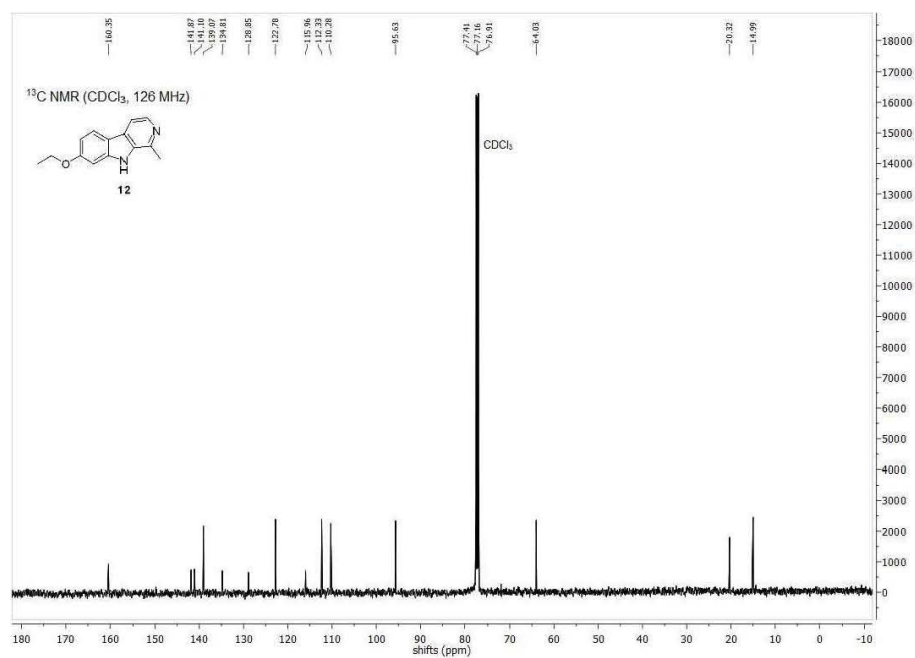
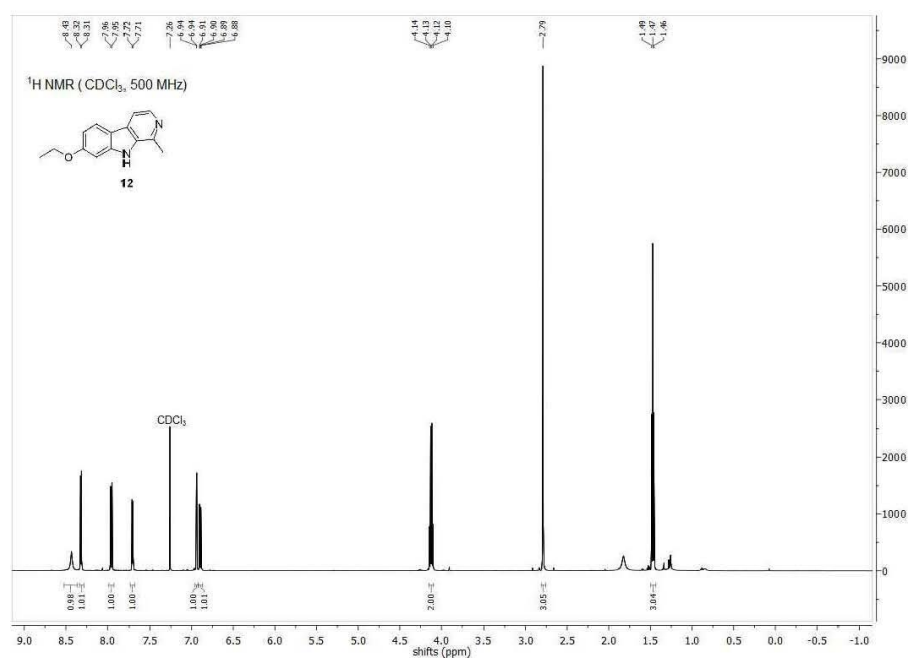


Compound 10





Compound 12

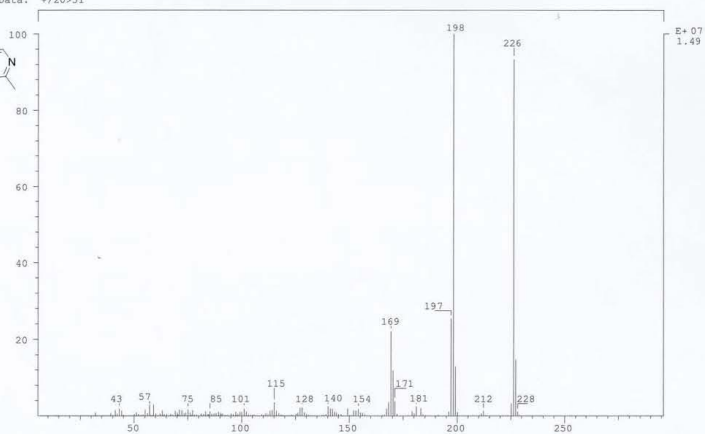
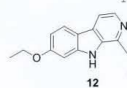


S27

MS (m/z, 70 eV, EI)

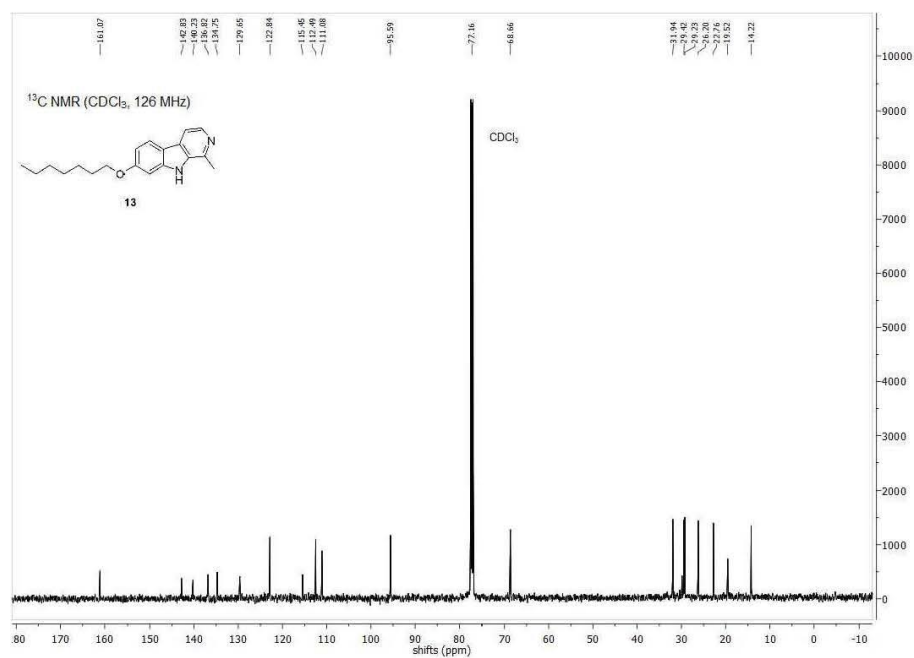
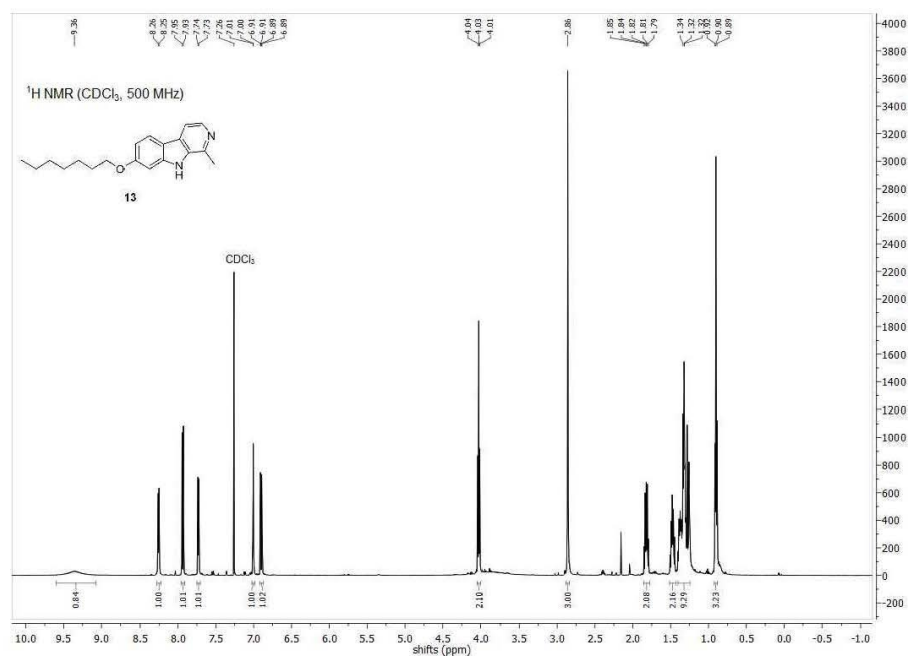
SPEC: 71165
Samp: BD 193 P2 A
Mode: EI +VE +LMR
Oper: Rudolph
Base: 198.2
Norm: 198.2
Peak: 1000.00
Data: +/-20x31

27-Jul-12 REG : 02:46.0 #9
Start : 11:37:27 31
BSCAN (EXP) UP LR NRM
Client: DRUNG/SCHMIDT
Inlet :
Inten : 14904623
Masses: 30 > 1000
#peaks: 539

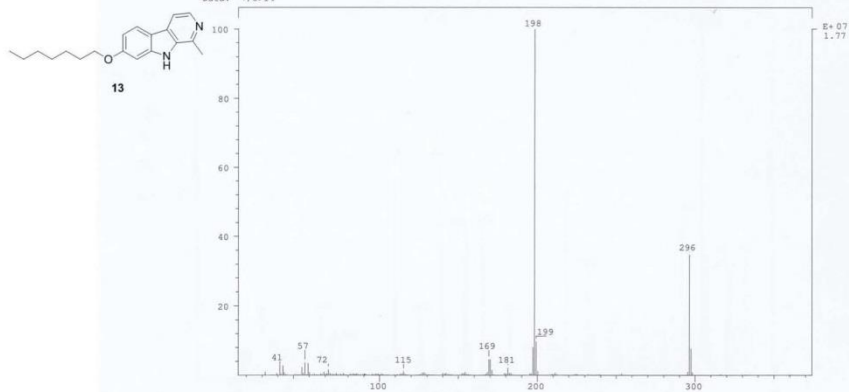


S28

Compound 13

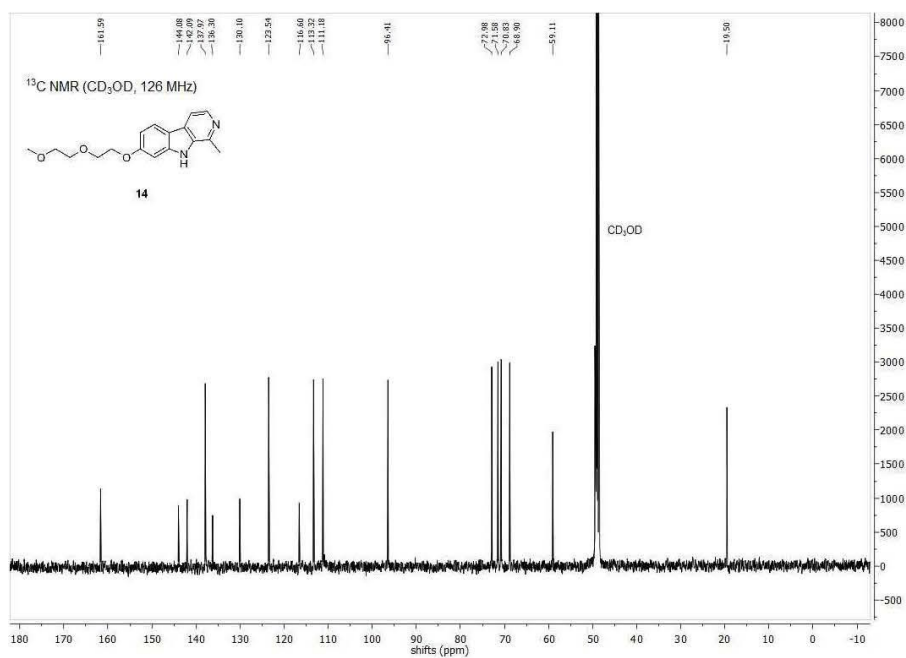
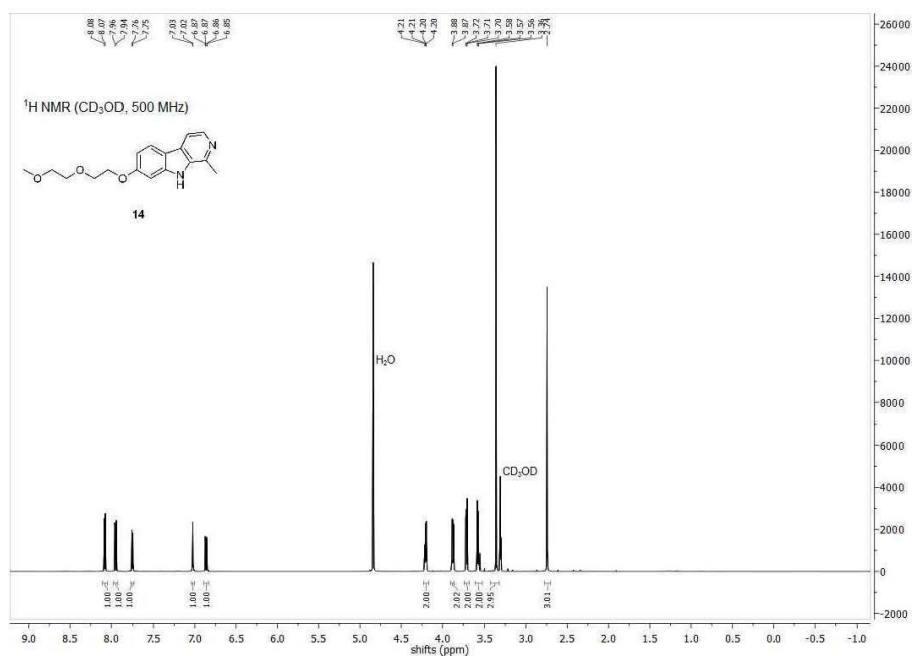


MS (m/z, 70 eV, EI)

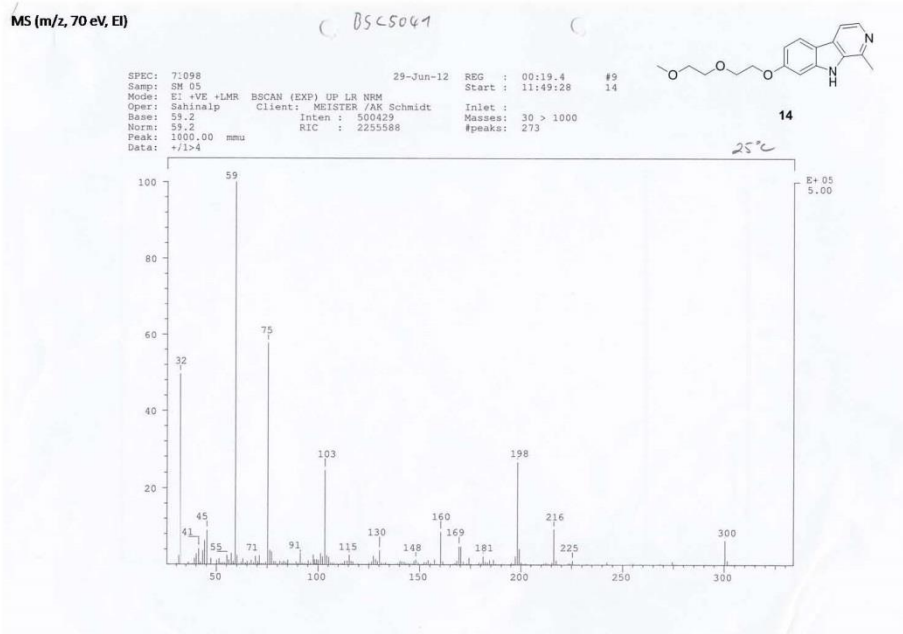


S30

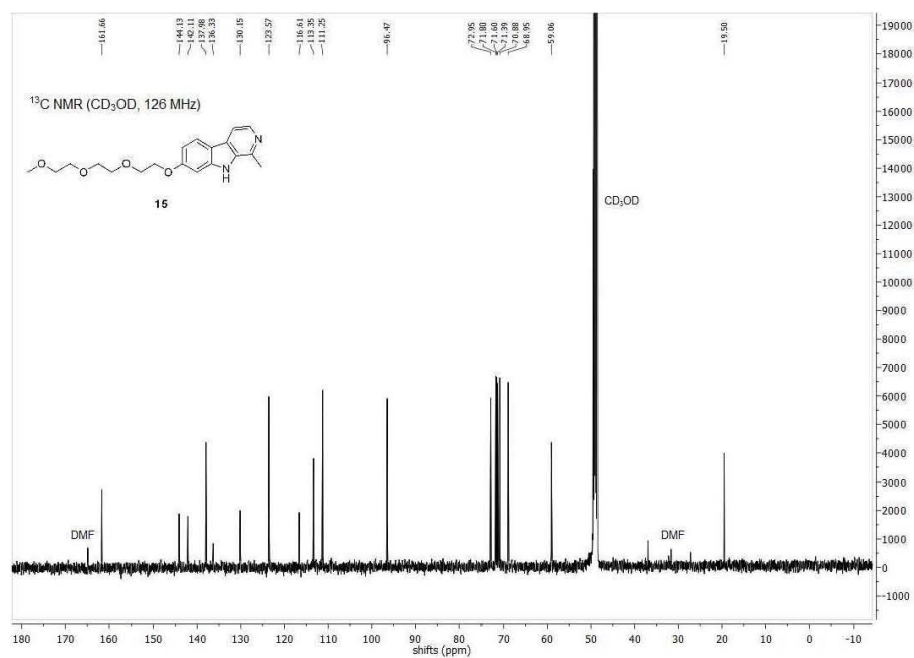
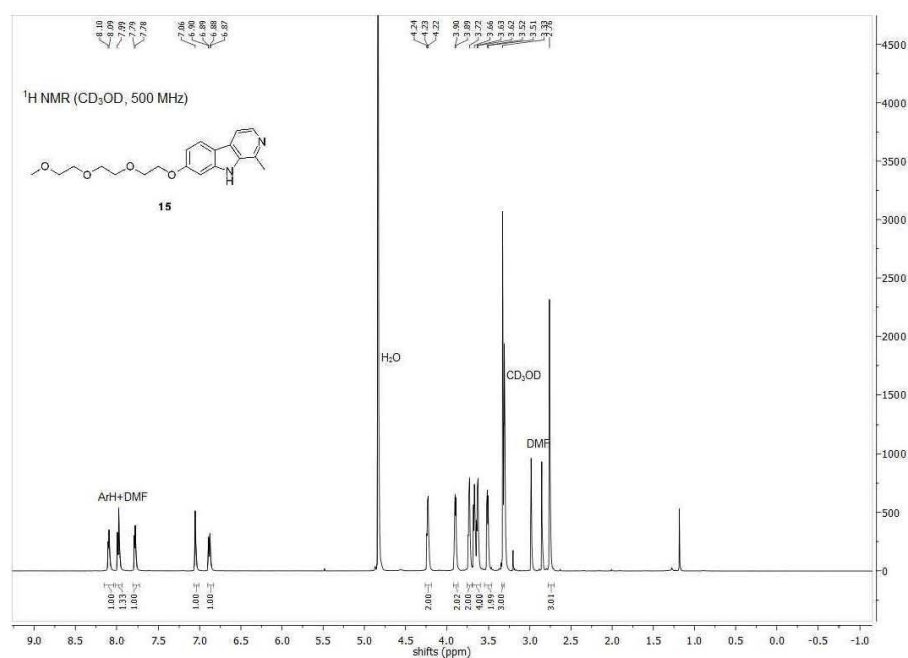
Compound 14



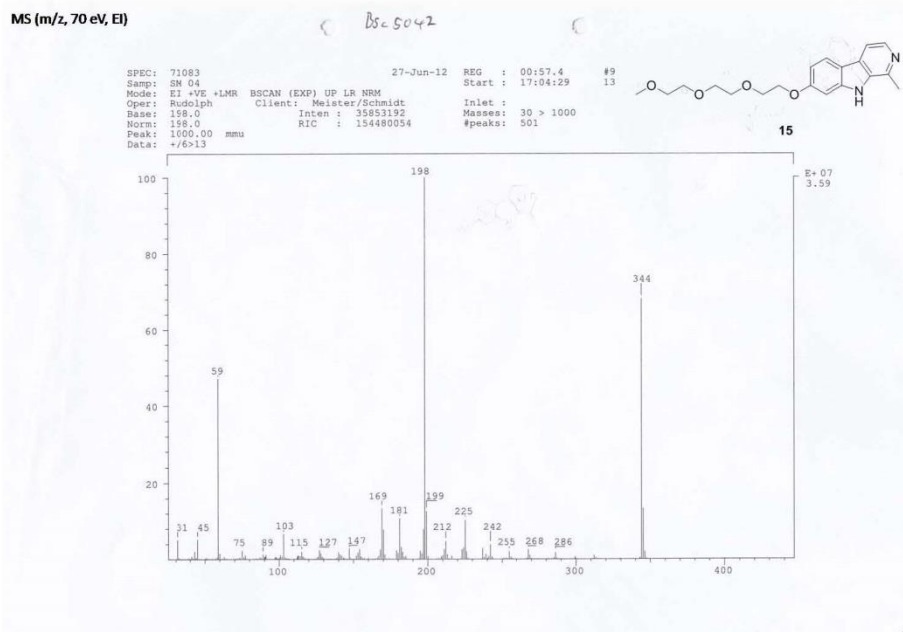
S31



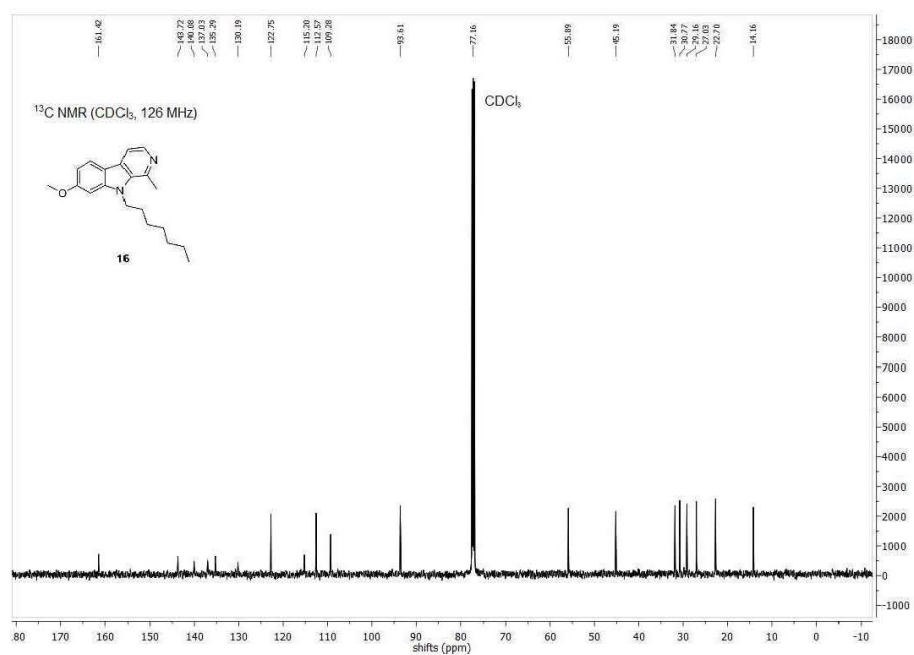
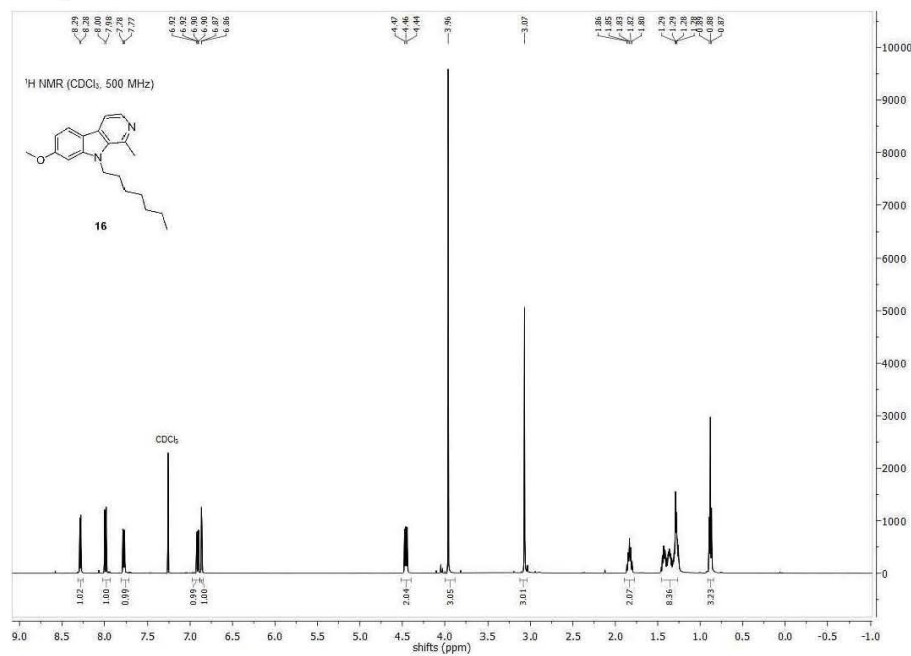
Compound 15



S33

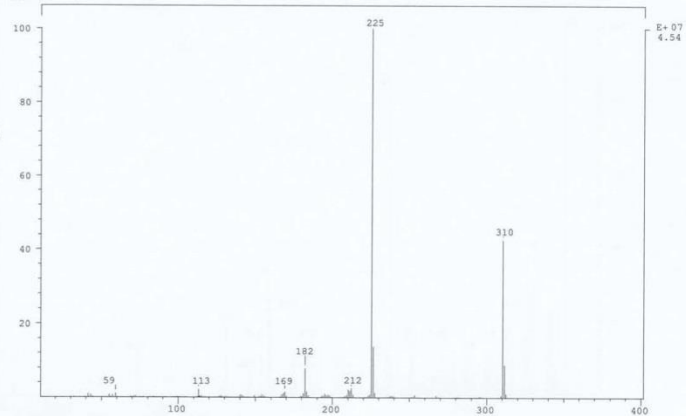
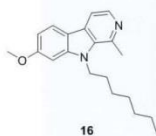


Compound 16

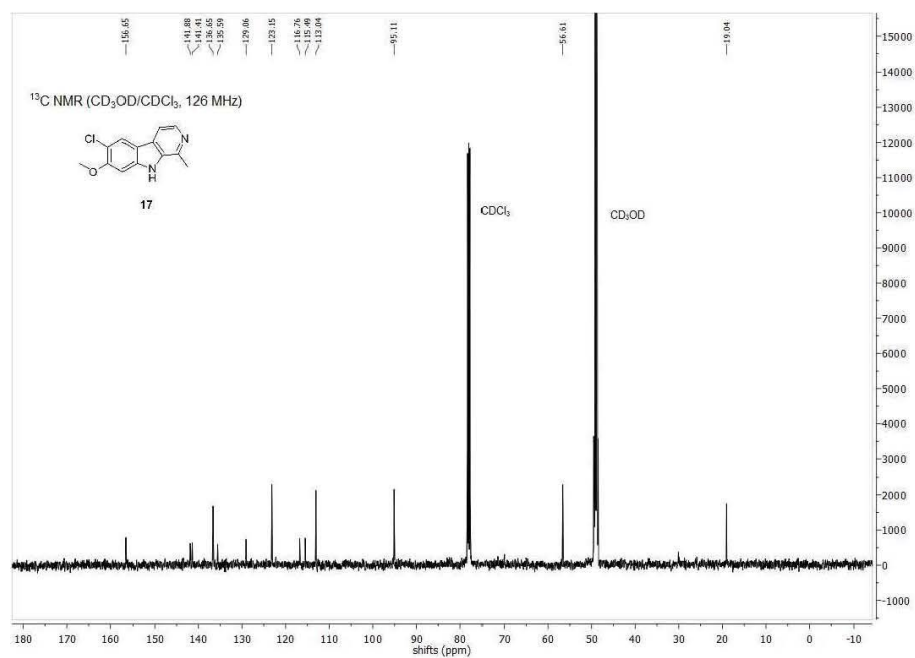
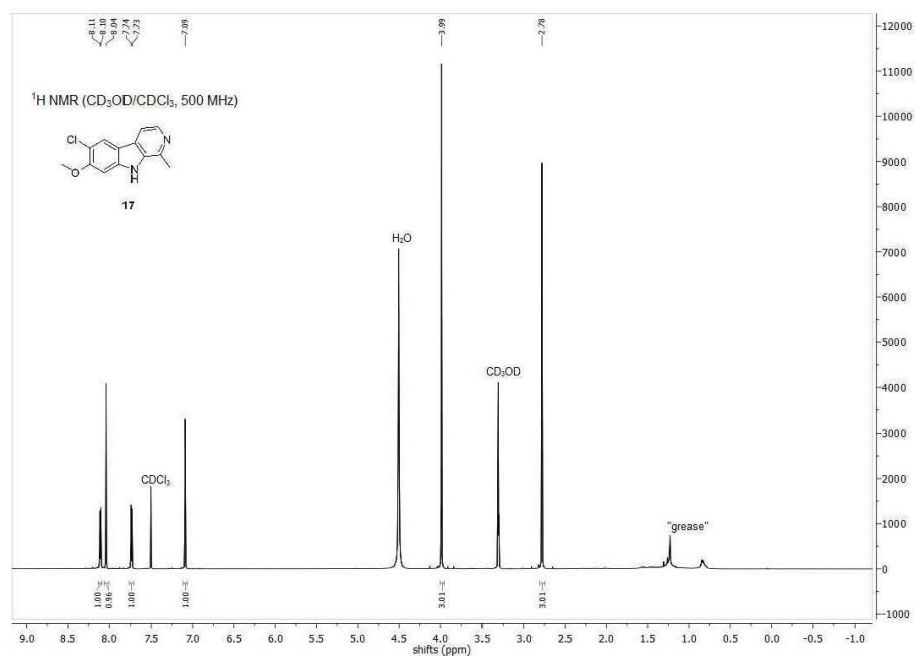


MS (m/z, 70 eV, EI)

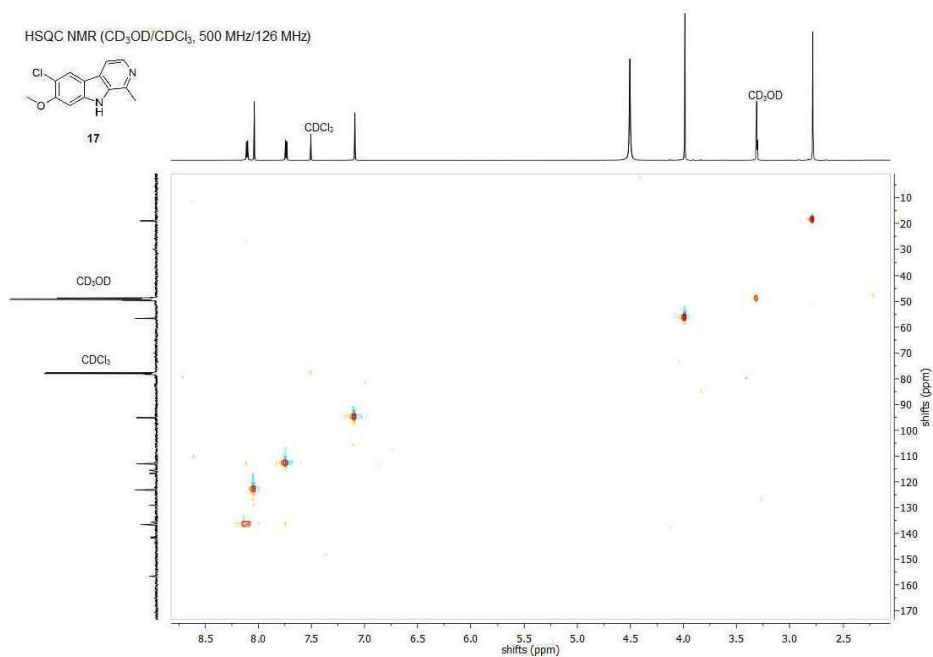
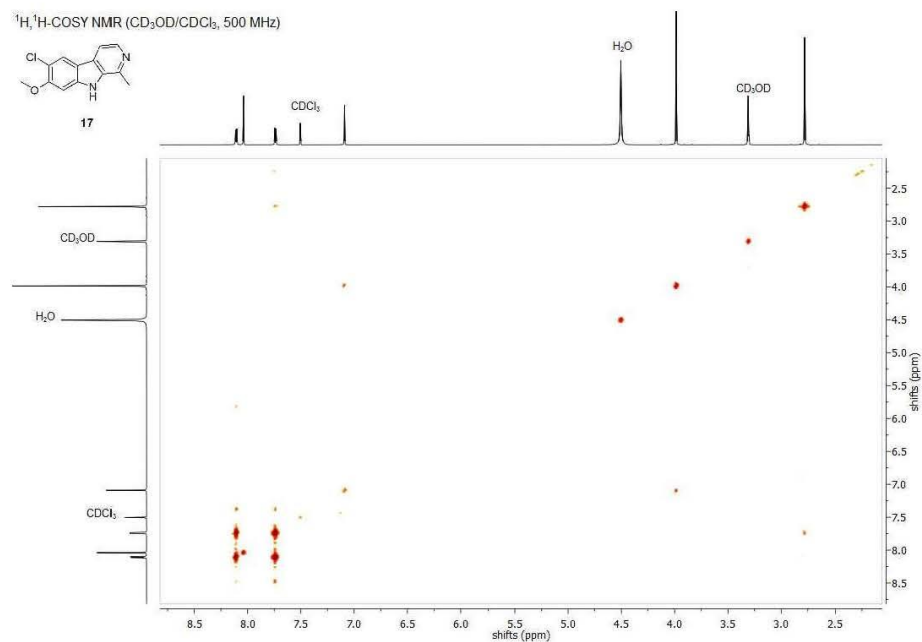
SPEC: 71167 27-Jul-12 RES : 01:09.0 #9
Samp: BD 191 P2 A Start : 11:59:55 17
Mode: EI +VE +LMR BSCAN (EXP) UP LR NRM
Oper: Rudolph Client: DRUGS/SCHMIDT Inlet :
Base: 225.2 Inten : 45388384 Masses: 30 > 1000
Norm: 225.2 RIC : 102563219 #peaks: 464
Peak: 1000.00 mmu
Data: +/-16



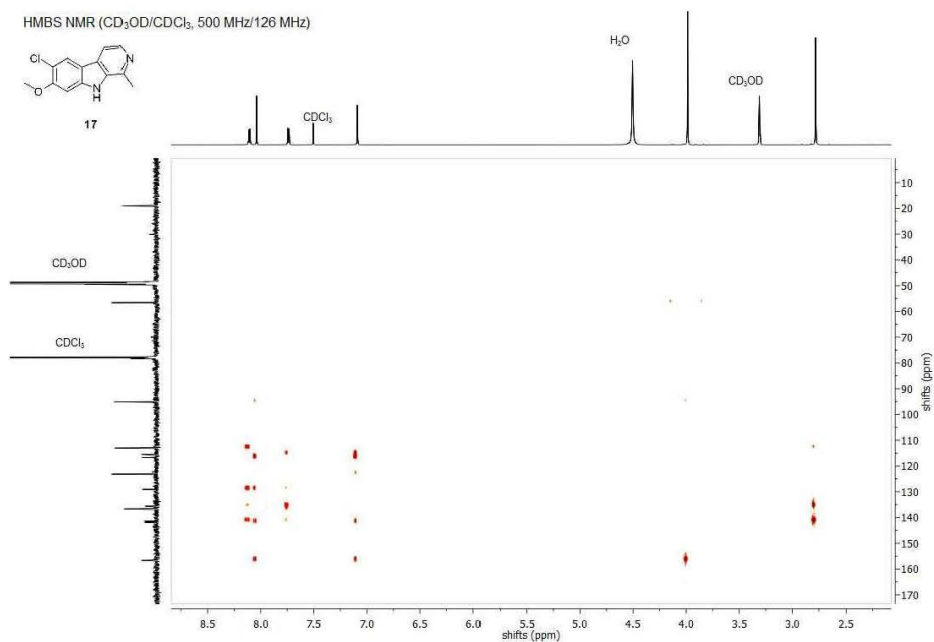
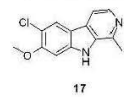
Compound 17



S37



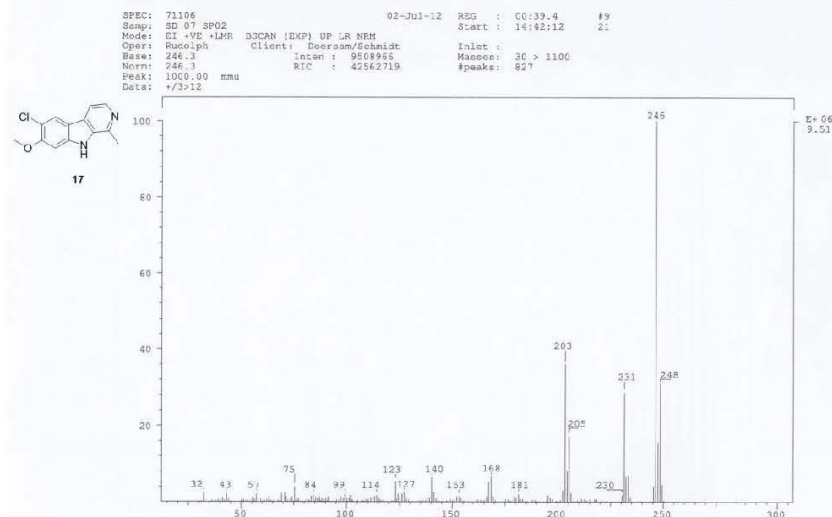
HMBS NMR (CD₃OD/CDCl₃, 500 MHz/126 MHz)



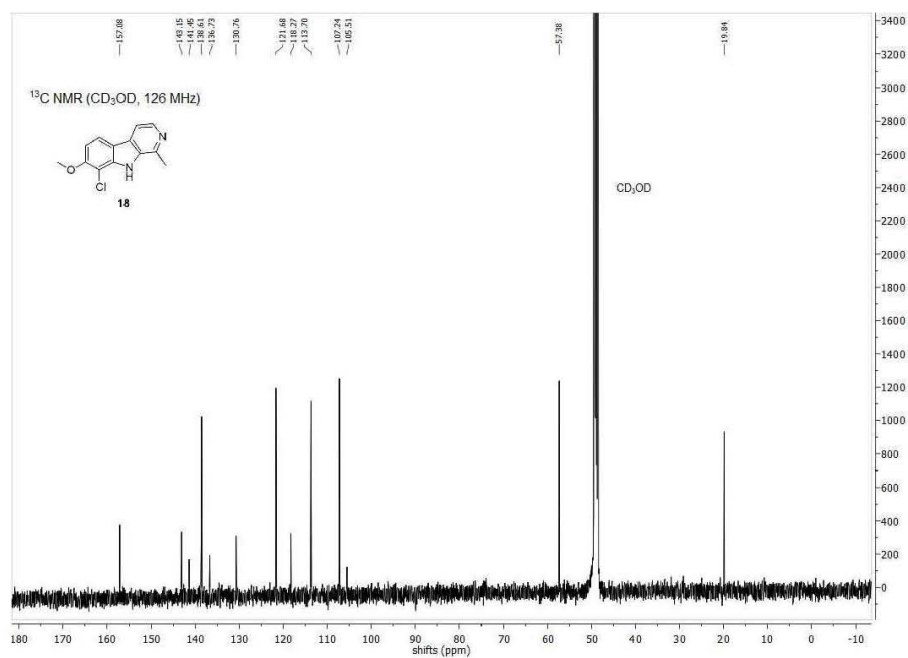
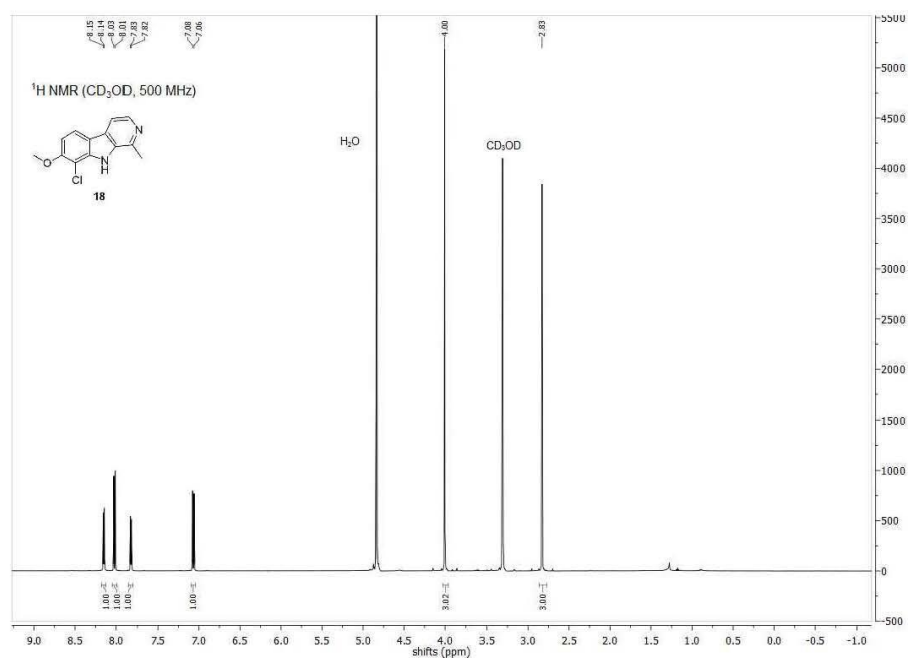
MS (m/z, 70 eV, EI)

DSC5049

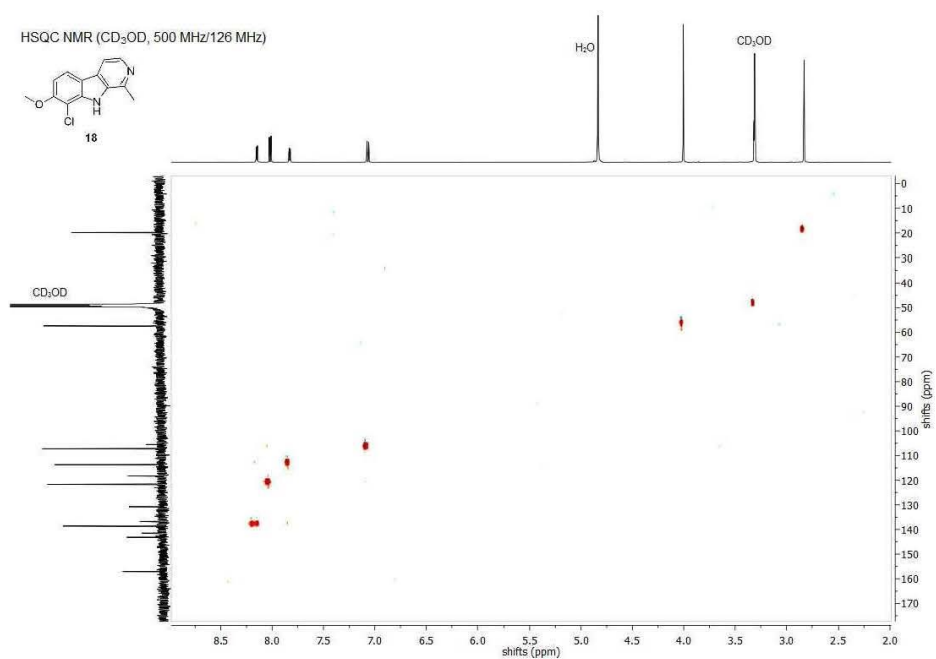
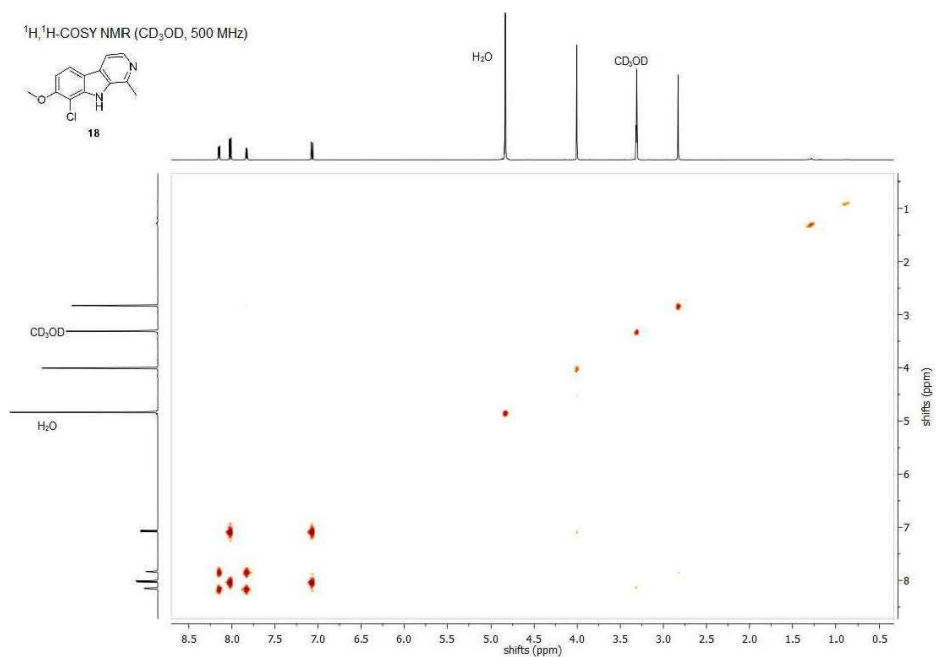
b11/16/07



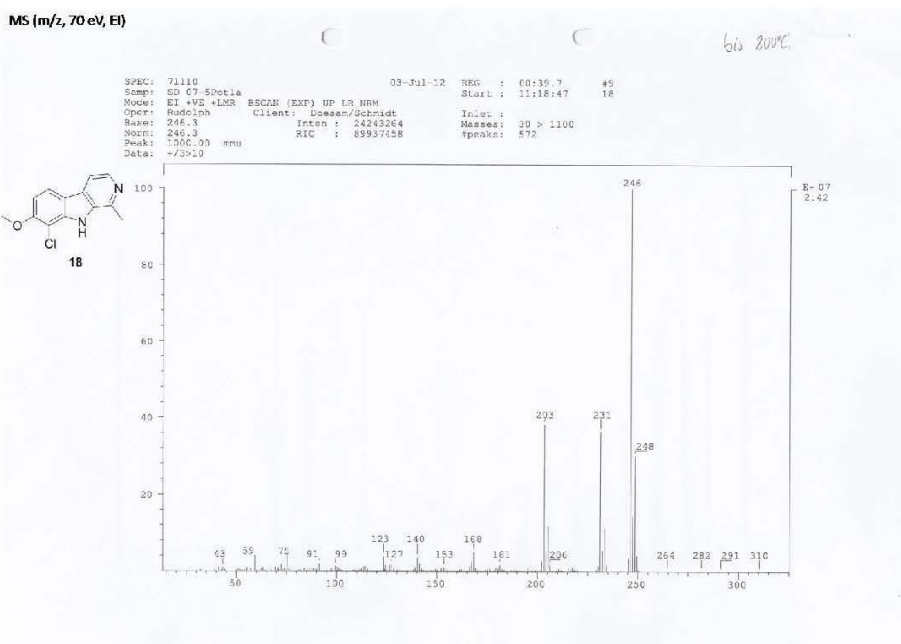
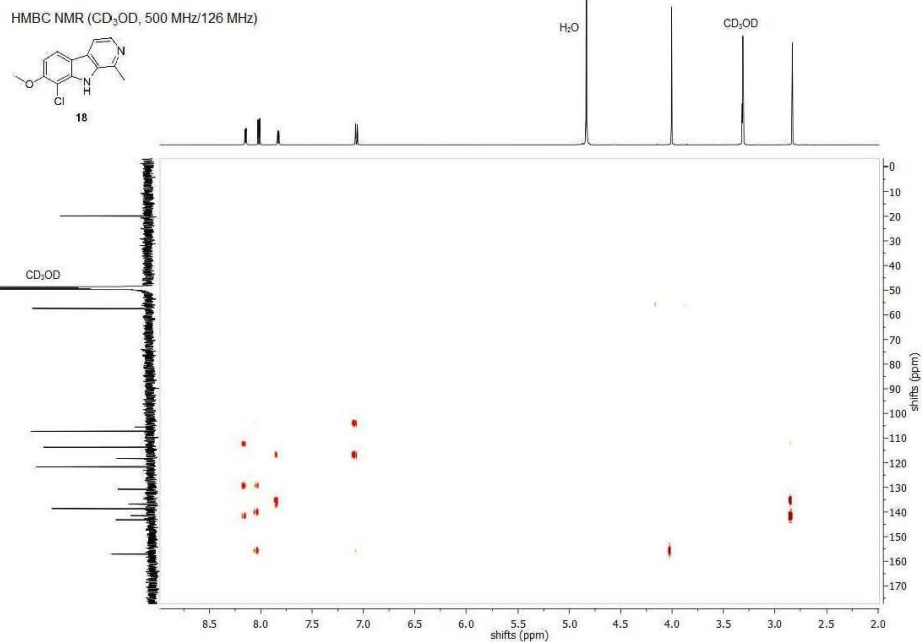
Compound 18



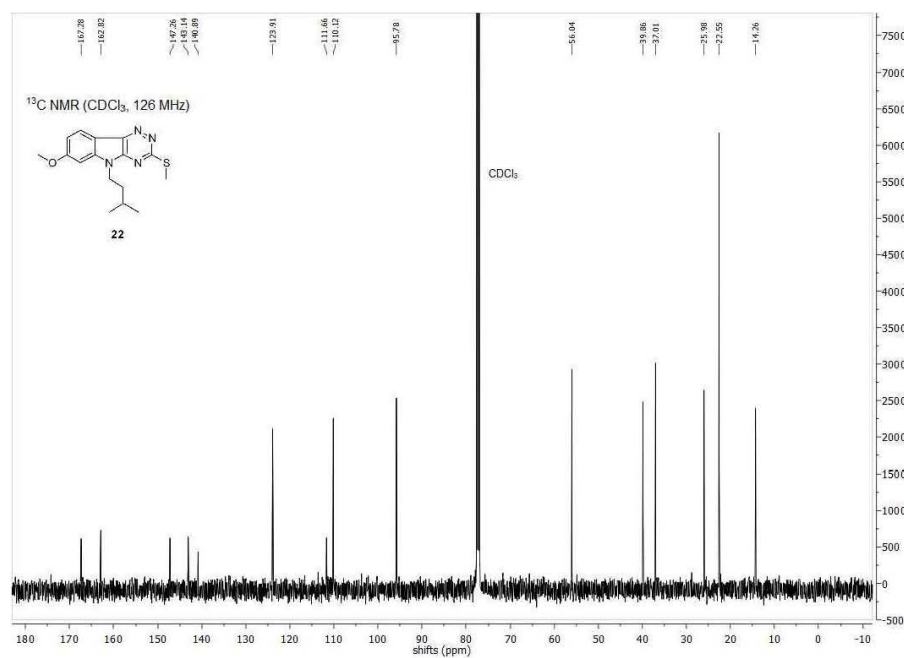
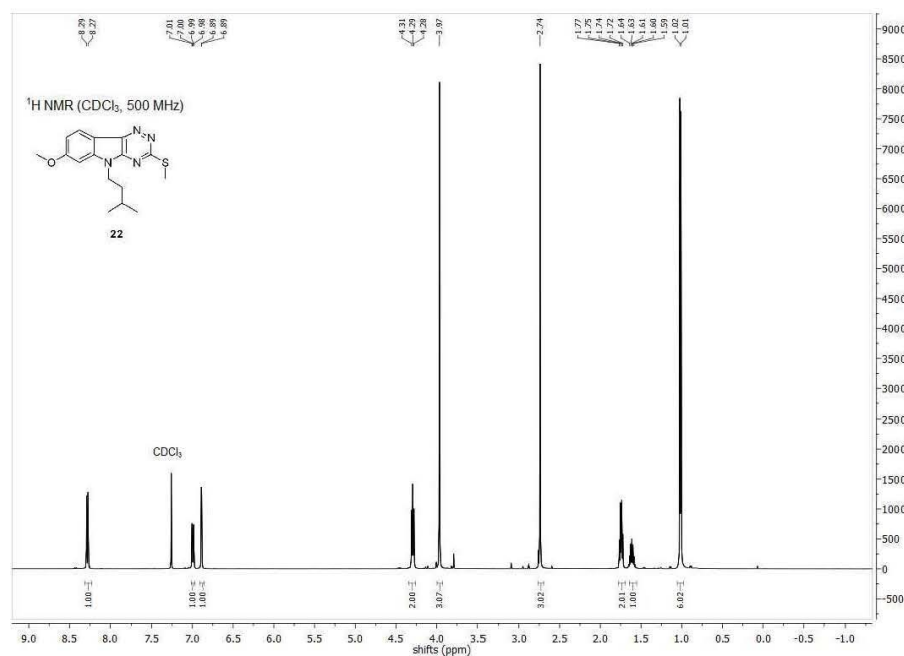
S40



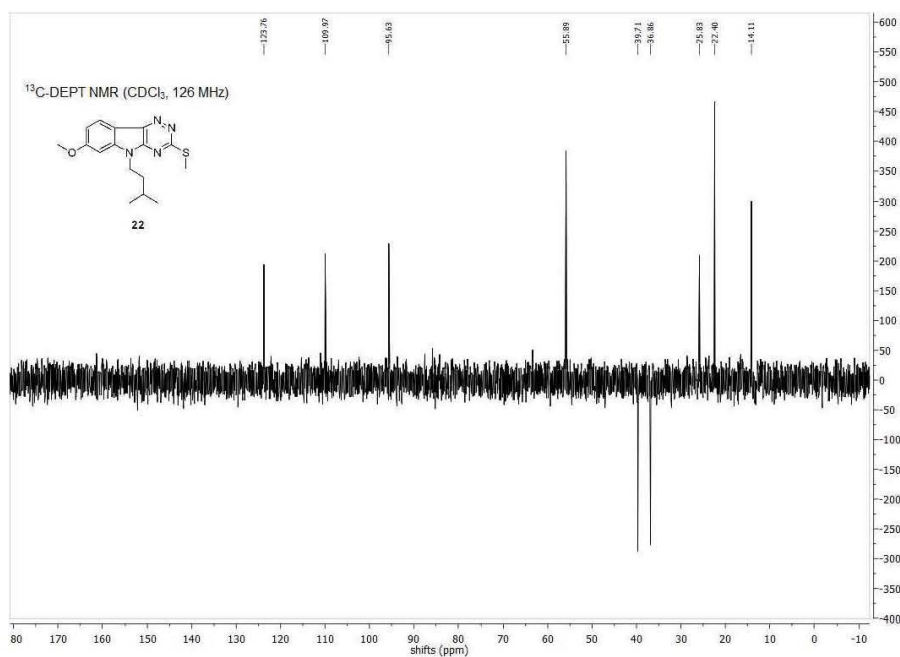
S41



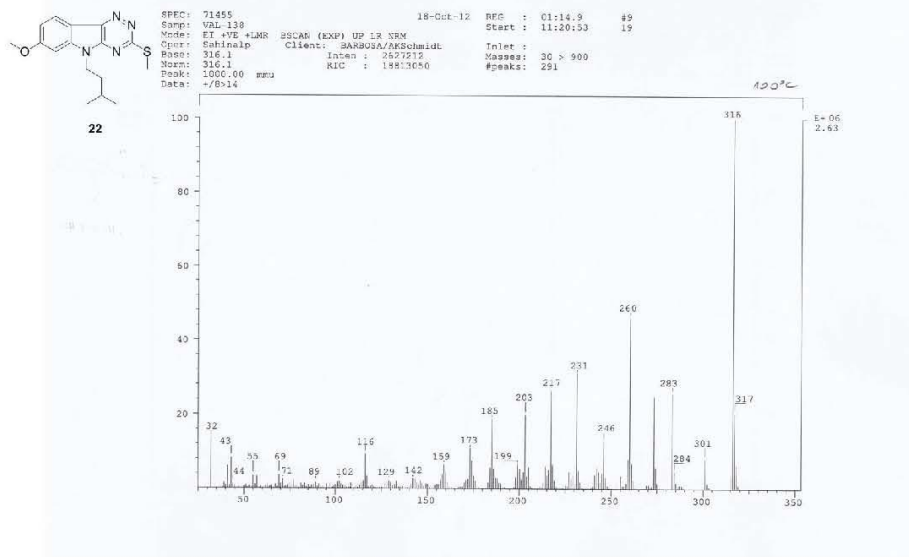
Compound 22



S43

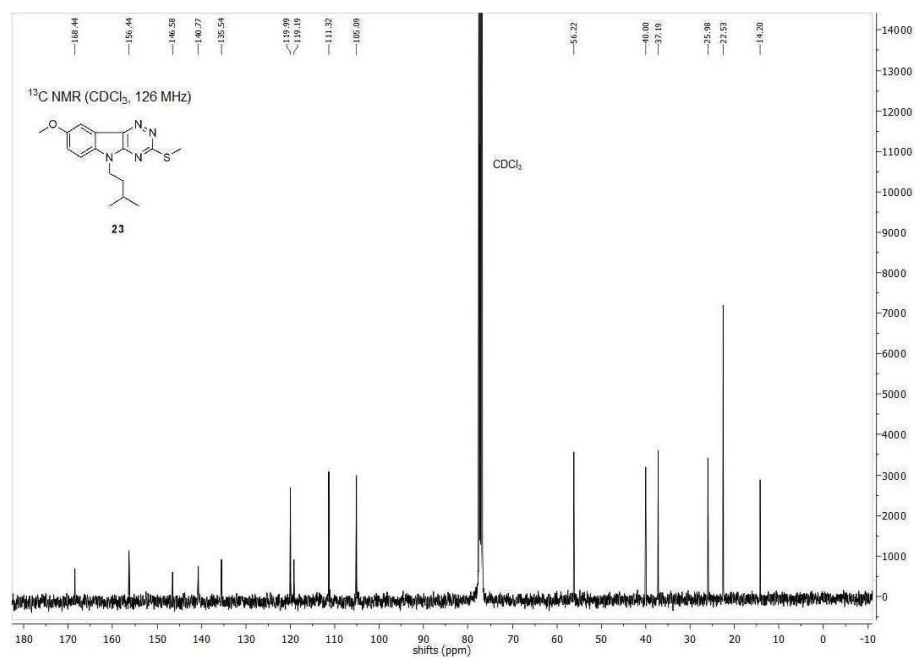
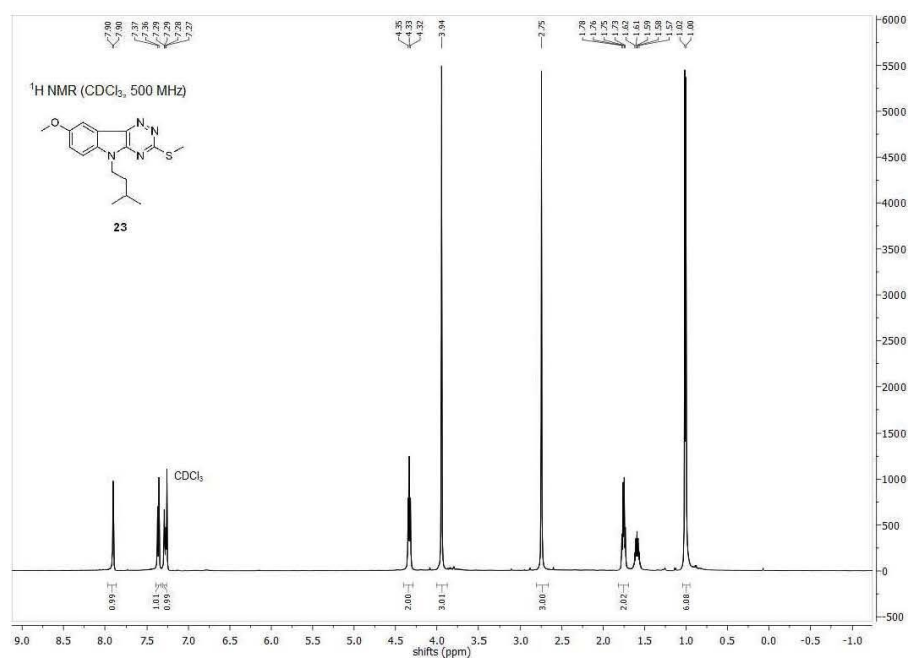


MS (m/z, 70 eV, EI)

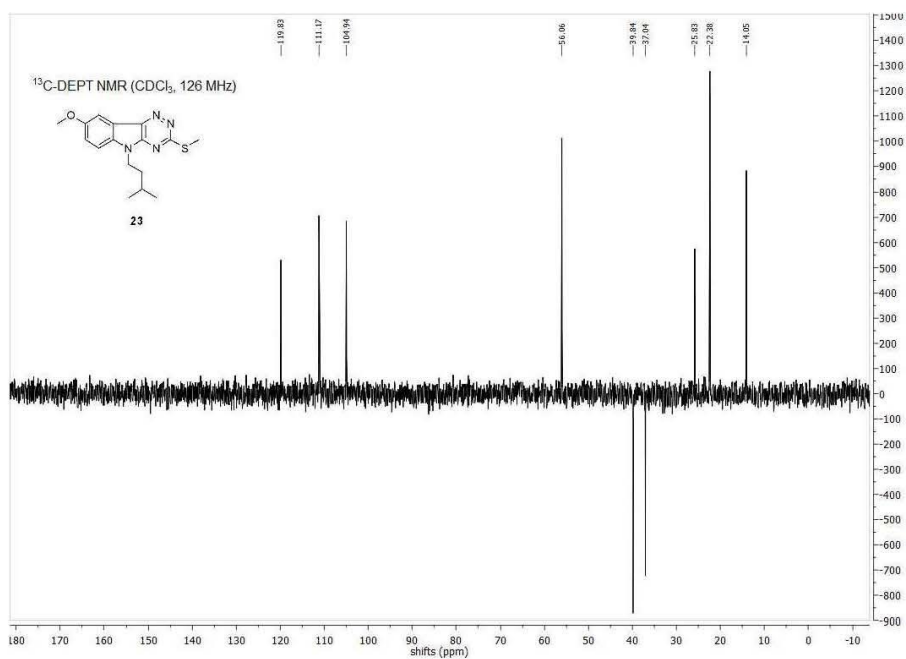


S44

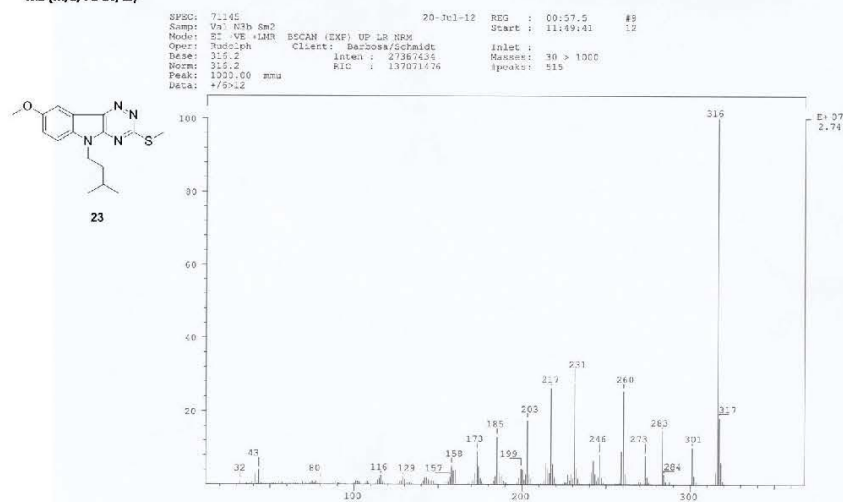
Compound 23



S45



MS (m/z, 70 eV, EI)



VII. References

- Mott, B. T.; Tanega, C.; Shen, M.; Maloney, D. J.; Shinn, P.; Leister, W.; Marugan, J. J.; Inglese, J.; Austin, C. P.; Misteli, T.; Auld, D. S.; Thomas, C. J. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6700.
- Rosenthal, A. S.; Tanega, C.; Shen, M.; Mott, B. T.; Bougie, J. M.; Nguyen, D.-T.; Misteli, T.; Auld, D. S.; Maloney, D. J.; Thomas, C. J. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3152.
- Cuny, G. D.; Ulyanova, N. P.; Patnaik, D.; Liu, J.-F.; Lin, X.; Auerbach, K.; Ray, S. S.; Xian, J.; Glicksman, M. A.; Stein, R. L.; Higgins, J. M. G. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2015.
- Bamford, M. J.; Bailey, N.; Davies, S.; Dean, D. K.; Francis, L.; Panchal, T. A.; Parr, C. A.; Sehmi, S.; Steadman, J. G.; Takle, A. K.; Townsend, J. T.; Wilson, D. M. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3407.
- Pagano, M. A.; Andrzejewska, M.; Ruzzene, M.; Sarno, S.; Cesaro, L.; Bain, J.; Elliott, M.; Meggio, F.; Kazimierzczuk, Z.; Pinna, L. A. *J. Med. Chem.* **2004**, *47*, 6239.
- J. Leonhard; B. Lygo; Procter, G. *Praxis in der organischen Chemie*. VCH Verlagsgesellschaft: Weinheim, 1996.
- Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512.
- (a) de Greef, T. F. A.; Nieuwenhuizen, M. M. L.; Sijbesma, R. P.; Meijer, E. W. *J. Org. Chem.* **2010**, *75*, 598; (b) Sun, J.; Stone, G. M.; Balsara, N. P.; Zuckermann, R. N. *Macromolecules* **2012**, *45*, 5151.
- Ponce, M. A.; Tarzi, O. I.; Erra-Balsells, R. *J. Heterocyclic Chem.* **2003**, *40*, 419.
- Mason, J. J.; Janosik, T.; Bergman, J. *Synthesis* **2009**, 2009, 3642.
- Himpel, S.; Tegge, W.; Frank, R.; Leder, S.; Joost, H.-G.; Becker, W. *J. Biol. Chem.* **2000**, *275*, 2431.
- Weyler, W.; Salach, J. I. *J. Biol. Chem.* **1985**, *260*, 13199.
- (a) Cao, R.; Peng, W.; Chen, H.; Hou, X.; Guan, H.; Chen, Q.; Ma, Y.; Xu, A. *Eur. J. Med. Chem.* **2005**, *40*, 249; (b) Zeng, Y.; Zhang, Y.; Weng, Q.; Hu, M.; Zhong, G. *Molecules* **2010**, *15*, 7775.
- Song, Y.; Wang, J.; Teng, S. F.; Kesuma, D.; Deng, Y.; Duan, J.; Wang, J. H.; Qi, R. Z.; Sim, M. M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1129.
- (a) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. *Nucleic Acids Res.* **2007**, *35*, D198; (b) Chen, X.; Lin, Y.; Liu, M.; Gilson, M. K. *Bioinformatics* **2002**, *18*, 130.
- Molecular Operating Environment (MOE)*, 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7: 2013.
- Halgren, T. A. *J. Comp. Chem.* **1996**, *17*, 490.
- Labute, P. *Proteins: Struct., Funct., Bioinf.* **2009**, *75*, 187.
- Labute, P. *J. Comput. Chem.* **2008**, *29*, 1693.
- (a) Gerber, P.; Müller, K. *J. Comput. Aided Mol. Des.* **1995**, *9*, 251; (b) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198.
- Neudert, G.; Klebe, G. *J. Chem. Inf. Model.* **2011**, *51*, 2731.
- <http://www.cerep.fr/cerep/users/pages/catalog/profiles/catalog.asp> (08/15/2014).
- Weyler, W.; Salach, J. I. *J. Biol. Chem.* **1985**, *260*, 13199.
- Weissbach, H.; Smith, T. E.; Daly, J. W.; Witkop, B.; Udenfriend, S. *J. Biol. Chem.* **1960**, *235*, 1160.
- Taira, N.; Nihira, K.; Yamaguchi, T.; Miki, Y.; Yoshida, K. *Mol. Cell* **25**, 725.
- Li, K.; Zhao, S.; Karur, V.; Wojchowski, D. M. *J. Biol. Chem.* **2002**, *277*, 47052.
- Muraki, M.; Ohkawara, B.; Hosoya, T.; Onogi, H.; Koizumi, J.; Koizumi, T.; Sumi, K.; Yomoda, J.-i.; Murray, M. V.; Kimura, H.; Furuichi, K.; Shibuya, H.; Krainer, A. R.; Suzuki, M.; Hagiwara, M. *J. Biol. Chem.* **2004**, *279*, 24246.
- Nayler, O.; Schnorrer, F.; Stamm, S.; Ullrich, A. *J. Biol. Chem.* **1998**, *273*, 34341.

4. Zusammenfassung und Ausblick

Übergeordnetes Ziel dieser Arbeit war die systematische Anwendung interdisziplinärer Methoden des modernen Wirkstoffdesigns sowie deren Weiter- und Neuentwicklung. Mit Bortezomib **5** und Carfilzomib **8** sind bereits zwei Inhibitoren des 20S-Proteasoms zur Behandlung schwerer Tumorerkrankungen des blutbildenden Systems zugelassen, die jedoch aufgrund ihres (pseudo)-irreversiblen Bindungsmodus nicht zur Behandlung solider Tumoren geeignet sind. α -Ketophenylamide, allen voran die Leitstruktur BSc2189 **9**, binden nicht nur nachweislich reversibel, sondern bieten durch den Kontakt mit der gestrichenen Seite der β 5-Untereinheit völlig neue Möglichkeiten, um Aktivität und Selektivität zu optimieren. Eine solche Optimierung ist mit dem in dieser Arbeit entwickelten Wirkstoff BSc4999 gelungen, der mit einem IC_{50} -Wert von 38 nM doppelt so aktiv ist wie die ursprüngliche Leitstruktur **9**. Die Verbindung ist hochselektiv für die β 5-Untereinheit und zudem zellgängig, was in diversen *in vitro*-Assays belegt werden konnte. Das Wirkstoffdesign erfolgte streng rational, wodurch der Zeit- und Materialaufwand mit nur fünf synthetisierten Derivaten minimal gehalten werden konnte. Die Analyse röntgenkristallographischer Aufnahmen war hierbei ebenso entscheidend wie die quantitative Struktur/Aktivitäts-Beziehung mit dem elektronischen Parameter σ und das kovalente Protein/Ligand-Docking zur Identifizierung des potentesten aus fünf möglichen Methyl-Derivaten.

Ein ebenso interdisziplinärer – wenn auch völlig andersartiger – Ansatz wurde zur Untersuchung des postulierten „Mausefallen-Mechanismus“ im aktiven Zentrum des Proteasoms verfolgt. Die Hypothese, dass es durch einen Ligand/Met45-Kontakt zu einer kinetischen Aktivierung der Substratspaltung in der β 5-Untereinheit des Proteasoms kommt, wurde zunächst mittels eines ligandenbasierten, chemisch-synthetischen Ansatzes und *in vitro*-Evaluation aldehydischer Inhibitoren untersucht. Diese ersten Ergebnisse und eine begleitend durchgeführte Moleküldynamiksimulation bestätigten zunächst die ursprüngliche „Mausefallen-Hypothese“ – durch eine anschließend durchgeführte detaillierte kinetische Analyse wurde sie jedoch zweifelsfrei widerlegt. Anhand der gesammelten Daten konnte dennoch eine erstaunliche Erkenntnis gewonnen werden: Die β 5-Untereinheit spaltet offensichtlich bevorzugt nach kleinen neutralen P1-Resten (SNAAP), während die in der Literatur zuerst oder ausschließlich genannte ChTL-Aktivität eine untergeordnete Rolle zu

spielen scheint. Dies hat direkte Einflüsse auf die Proteasomforschung, da somit neben einer Reklassifizierung in Elastase-ähnlich (EL) auch die Verwendung von Referenzsubstraten mit kleinen neutralen P1-Resten indiziert ist, um einen hohen Substratumsatz mit maximaler β 5-Selektivität zu gewährleisten.

Die überwiegende Mehrheit aller bisher entwickelten Proteasominhibitoren binden kovalent an das aktive Zentrum. Kovalente Inhibitoren haben oftmals wesentliche Vorteile, wie beispielsweise eine höhere biologische Aktivität, Selektivität, eine geringere Dosierung, weniger häufige Applikation des Wirkstoffs bei der Behandlung und bessere pharmakokinetische Eigenschaften. Dies lässt sie mittlerweile auch für die Forschung an anderen Zielproteinen und für die Pharmaindustrie zunehmend interessant erscheinen. Für ein rationales strukturbasiertes Design kovalenter Inhibitoren ist das Protein/Ligand-Docking eine wertvolle Methode, wobei die verfügbaren Softwarelösungen in Bezug auf Benutzerfreundlichkeit, Vielseitigkeit und Schnelligkeit erhöhten Optimierungsbedarf zeigen. Mit der Software DOCKTITE konnte im Rahmen dieser Arbeit eine Methode zum kovalenten Docking entwickelt werden, die keinerlei Restriktionen bezüglich elektrophiler Kopfgruppen, Proteinklassen oder nukleophiler Seitenketten zeigt und deren Präzision und Schnelligkeit in einer umfangreichen Validierungsstudie belegt wurde. Der modulare Aufbau und der frei verfügbare Quellcode von DOCKTITE ermöglichen erfahrenen Benutzern eigene Bindungsmechanismen äußerst komfortabel über eine 2D-Zeichensoftware zuzufügen, einzelne Schritte in andere Programme auszulagern oder die SVL-Skripte je nach Bedarf anzupassen. Einige Funktionen von DOCKTITE wurden basierend auf regem und überaus erfreulichem wissenschaftlichem Austausch mit Anwendern der Merck KGaA implementiert. Diese inspirierende Zusammenarbeit hat DOCKTITE den Einzug in ein international tätiges Pharmaunternehmen ermöglicht und wird sicher auch in Zukunft fortbestehen.

Neben der Entwicklung von Antitumorwirkstoffen befanden sich die ALZHEIMER-Demenz und die Entwicklung potentieller diagnostischer Sonden und Wirkstoffe im Fokus dieser Arbeit. Es konnte eine neue Syntheseroute für Derivate des etablierten Fluoreszenzliganden zur Markierung von amyloiden Plaques und Neurofibrillenbündeln Methoxy-XO4 **13** entwickelt werden. Über diese optimierte Syntheseroute wurden elf neue Derivate der Leitstruktur **13** mit geringem Aufwand synthetisiert und histologisch sowie in einem Verdrängungs-Assay untersucht. Dabei wurden erstmals

Methoxy-XO4-Derivate mit erhöhter Selektivität gegenüber jeder der genannten pathogenen Ablagerungen identifiziert, was der Fluoreszenzmikroskopie als bildgebendes Verfahren in der Diagnostik von Morbus ALZHEIMER völlig neue Einsatzmöglichkeiten eröffnet. Durch die selektive Detektion von NFTs und APs in der Retina – unabhängig voneinander – ist es nun möglich, frühzeitig eine nötige Therapie einzuleiten bzw. den Krankheitsverlauf und den Therapieerfolg optimal zu überwachen.

Die Kinase DYRK1A ist ein Zielprotein (*Target*), das zum einen relativ wenig erforscht ist und zum anderen mit beiden Theorien zur Entstehung der ALZHEIMER-Demenz in Verbindung steht: Der Amyloid- und der Tau-Hypothese. Bei einem derart neuen *Target* ist es zunächst wichtig die Eigenschaften eines Wirkstoffs zu identifizieren, die für dessen pharmakologische Wirkung essentiell sind, also seine Struktur/Wirkungs-Beziehung aufzuklären. Ein solcher Ausgangspunkt stand mit dem β -Carbolin Harmin **14** zur Verfügung, welches jedoch durch Ko-Inhibition von MAO-A zu starken Nebenwirkungen *in vivo* führt. Durch eine ligandenbasierte Pharmakophor-Modellierung mit einer Datenbank aus 127 Verbindungen konnten zunächst wertvolle Informationen bezüglich dieser essentiellen Eigenschaften gewonnen werden. Durch Protein/Ligand-Docking und Synthese diverser Harmin-Derivate konnte das computerbasierte Modell bestätigt und verfeinert werden. Anschließend wurde in einem strukturbasierten Ansatz mittels Docking, Analyse wissensbasierter statistischer Potentiale, chemischer Synthese und biochemischer *in vitro*-Evaluation ein potenter DYRK1A-Inhibitor mit signifikanter Selektivität gegenüber DYRK2, DYRK3, DYRK4, CLK2 und natürlich MAO-A entwickelt. Die in dieser Arbeit gewonnen Erkenntnisse erleichtern das zukünftige Design neuer DYRK1A-Inhibitoren und geben zweifellos neue Impulse, die den β -Carbolinen in den nächsten Jahren Einzug in klinische Studien für diese innovative Therapie der ALZHEIMER-Demenz ermöglichen könnten.

- [1] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, "The Protein Data Bank", *Nucleic Acids Res.* **2000**, *28*, 235-242.
- [2] R. Schoenheimer, *The Dynamic State of Body Constituents*, Harvard University Press, Cambridge, **1942**.
- [3] M. V. Simpson, "The release of labeled amino acids from the proteins of rat liver slices", *J. Biol. Chem.* **1953**, *201*, 143-154.
- [4] C. D. Duve, R. Gianetto, F. Appelmans, R. Wattiaux, "Enzymic Content of the Mitochondria Fraction", *Nature* **1953**, *172*, 1143-1144.
- [5] A. Ciechanover, "Intracellular protein degradation: From a vague idea thru the lysosome and the ubiquitin–proteasome system and onto human diseases and drug targeting", *BBA - Proteins Proteom.* **2012**, *1824*, 3-13.
- [6] http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2004/ (Zugriff: 26.03.2015).
- [7] M. Biel, V. Wascholowski, A. Giannis, "Eine verhängnisvolle Affäre: die Ubiquitylierung von Proteinen", *Angew. Chem.* **2004**, *116*, 6574-6576.
- [8] A. Ehlinger, K. J. Walters, "Structural Insights into Proteasome Activation by the 19S Regulatory Particle", *Biochemistry* **2013**, *52*, 3618-3628.
- [9] L. Borissenko, M. Groll, "20S Proteasome and its Inhibitors: Crystallographic Knowledge for Drug Development", *Chem. Rev.* **2007**, *107*, 687-717.
- [10] C. Voss, Dissertation, TU Darmstadt (Darmstadt, Germany), **2014**.
- [11] A. F. Kisselev, W. A. van der Linden, H. S. Overkleeft, "Proteasome Inhibitors: An Expanding Army Attacking a Unique Target", *Chem. Biol.* **2012**, *19*, 99-115.
- [12] I. Schechter, A. Berger, "On the size of the active site in proteases. I. Papain", *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157-162.
- [13] M. Groll, W. Heinemeyer, S. Jäger, T. Ullrich, M. Bochtler, D. H. Wolf, R. Huber, "The catalytic sites of 20S proteasomes and their role in subunit maturation: A mutational and crystallographic study", *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10976-10983.
- [14] T. P. Dick, A. K. Nussbaum, M. Deeg, W. Heinemeyer, M. Groll, M. Schirle, W. Keilholz, S. Stevanović, D. H. Wolf, R. Huber, H.-G. Rammensee, H. Schild, "Contribution of Proteasomal β -Subunits to the Cleavage of Peptide Substrates Analyzed with Yeast Mutants", *J. Biol. Chem.* **1998**, *273*, 25637-25646.
- [15] A. K. Nussbaum, T. P. Dick, W. Keilholz, M. Schirle, S. Stevanović, K. Dietz, W. Heinemeyer, M. Groll, D. H. Wolf, R. Huber, H.-G. Rammensee, H. Schild, "Cleavage

- motifs of the yeast 20S proteasome β subunits deduced from digests of enolase 1", *Proc. Natl. Acad. Sci. USA* **1998**, 95, 12504-12509.
- [16] Eva M. Huber, M. Basler, R. Schwab, W. Heinemeyer, Christopher J. Kirk, M. Groettrup, M. Groll, "Immuno- and Constitutive Proteasome Crystal Structures Reveal Differences in Substrate and Inhibitor Specificity", *Cell* **2012**, 148, 727-738.
- [17] B. Alberts, *Molecular Biology of the Cell: Reference Edition*, Garland Science, **2008**.
- [18] D. P. Lane, "p53, guardian of the genome", *Nature* **1992**, 358, 15-16.
- [19] J. Lowe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, R. Huber, "Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution", *Science* **1995**, 268, 533-539.
- [20] E. Seemuller, A. Lupas, D. Stock, J. Lowe, R. Huber, W. Baumeister, "Proteasome from *Thermoplasma acidophilum*: a threonine protease", *Science* **1995**, 268, 579-582.
- [21] H. A. Braun, S. Umbreen, M. Groll, U. Kuckelkorn, I. Mlynarczuk, M. E. Wigand, I. Drung, P.-M. Kloetzel, B. Schmidt, "Tripeptide Mimetics Inhibit the 20 S Proteasome by Covalent Bonding to the Active Threonines", *J. Biol. Chem.* **2005**, 280, 28394-28401.
- [22] J. Adams, M. Behnke, S. Chen, A. A. Cruickshank, L. R. Dick, L. Grenier, J. M. Klunder, Y.-T. Ma, L. Plamondon, R. L. Stein, "Potent and selective inhibitors of the proteasome: Dipeptidyl boronic acids", *Bioorg. Med. Chem. Lett.* **1998**, 8, 333-338.
- [23] M. L. Stein, H. Cui, P. Beck, C. Dubiella, C. Voss, A. Krüger, B. Schmidt, M. Groll, "Systematic Comparison of Peptidic Proteasome Inhibitors Highlights the α -Ketoamide Electrophile as an Auspicious Reversible Lead Motif", *Angew. Chem. Int. Ed.* **2014**, 53, 1679-1683.
- [24] A. Appel, "Drugs: More shots on target", *Nature* **2011**, 480, S40-S42.
- [25] M. Iqbal, P. A. Messina McLaughlin, D. Dunn, S. Mallya, J. Husten, M. A. Ator, S. Chatterjee, "Proteasome inhibitors for cancer therapy", *Bioorg. Med. Chem.* **2012**, 20, 2362-2368.
- [26] M. Groll, C. R. Berkers, H. L. Ploegh, H. Ova, "Crystal Structure of the Boronic Acid-Based Proteasome Inhibitor Bortezomib in Complex with the Yeast 20S Proteasome", *Structure* **2006**, 14, 451-456.
- [27] S. Arastu-Kapur, J. L. Anderl, M. Kraus, F. Parlati, K. D. Shenk, S. J. Lee, T. Muchamuel, M. K. Bennett, C. Driessen, A. J. Ball, C. J. Kirk, "Nonproteasomal Targets of the Proteasome Inhibitors Bortezomib and Carfilzomib: a Link to Clinical Adverse Events", *Clin. Cancer Res.* **2011**, 17, 2734-2743.
- [28] L. M. Martins, A. Morrison, K. Klupsch, V. Fedele, N. Moiso, P. Teismann, A. Abuin, E. Grau, M. Geppert, G. P. Livi, C. L. Creasy, A. Martin, I. Hargreaves, S. J. Heales, H.

- Okada, S. Brandner, J. B. Schulz, T. Mak, J. Downward, "Neuroprotective Role of the Reaper-Related Serine Protease HtrA2/Omi Revealed by Targeted Deletion in Mice", *Mol. Cell. Biol.* **2004**, *24*, 9848-9862.
- [29] D. Graham-Rowe, "Overview: Multiple lines of attack", *Nature* **2011**, *480*, S34-S35.
- [30] Eva M. Huber, W. Heinemeyer, M. Groll, "Bortezomib-Resistant Mutant Proteasomes: Structural and Biochemical Evaluation with Carfilzomib and ONX 0914", *Structure*, *23*, 407-417.
- [31] D. S. Siegel, "From clinical trials to clinical practice: single-agent carfilzomib adverse events and their management in patients with relapsed and/or refractory multiple myeloma", *Ther. Adv. Hematol.* **2013**, *4*, 354-365.
- [32] D. S. Siegel, T. Martin, M. Wang, R. Vij, A. J. Jakubowiak, S. Lonial, S. Trudel, V. Kukreti, N. Bahlis, M. Alsina, A. Chanan-Khan, F. Buadi, F. J. Reu, G. Somlo, J. Zonder, K. Song, A. K. Stewart, E. Stadtmauer, L. Kunkel, S. Wear, A. F. Wong, R. Z. Orlowski, S. Jagannath, "A phase 2 study of single-agent carfilzomib (PX-171-003-A1) in patients with relapsed and refractory multiple myeloma", *Blood* **2012**, *120*, 2817-2825.
- [33] O. A. O'Connor, A. K. Stewart, M. Vallone, C. J. Molineaux, L. A. Kunkel, J. F. Gerecitano, R. Z. Orlowski, "A Phase 1 Dose Escalation Study of the Safety and Pharmacokinetics of the Novel Proteasome Inhibitor Carfilzomib (PR-171) in Patients with Hematologic Malignancies", *Clin. Cancer Res.* **2009**, *15*, 7085-7091.
- [34] S. Chatterjee, D. Dunn, S. Mallya, M. A. Ator, "P'-extended α -ketoamide inhibitors of proteasome", *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2603-2606.
- [35] M. J. Williamson, M. D. Silva, J. Terkelsen, R. Robertson, L. Yu, C. Xia, P. Hatsis, B. Bannerman, T. Babcock, Y. Cao, E. Kupperman, "The relationship among tumor architecture, pharmacokinetics, pharmacodynamics, and efficacy of bortezomib in mouse xenograft models", *Mol. Cancer Ther.* **2009**, *8*, 3234-3243.
- [36] J. H. v. t. Hoff, *Die Lagerung der Atome im Raume (aus dem Französischen "La chimie dans l'espace", erschienen 1875)*, Friedrich Vieweg und Sohn, Braunschweig, **1877**.
- [37] P. J. Ramberg, *Chemical structure, spatial arrangement : the early history of stereochemistry, 1874-1914*, Ashgate, Aldershot, Hampshire u.a., **2003**.
- [38] H. Kolbe, "Zeichen der Zeit II", *J. Prakt. Chem.* **1877**, *15*, 473-477.
- [39] J. D. Watson, F. H. C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid", *Nature* **1953**, *171*, 737-738.
- [40] G. Klebe, *Wirkstoffdesign*, 2nd ed., Spektrum Akademischer Verlag, Heidelberg, **2009**.
- [41] Z. Kolská, M. Zábranský, A. Randová, *Group Contribution Methods for Estimation of Selected Physico-Chemical Properties of Organic Compounds*, **2012**.

- [42] <http://www.cas.org/about-cas/cas-fact-sheets/registry-fact-sheet> (Zugriff: 31.03.2015).
- [43] T. J. Hou, K. Xia, W. Zhang, X. J. Xu, "ADME Evaluation in Drug Discovery. 4. Prediction of Aqueous Solubility Based on Atom Contribution Approach", *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 266-275.
- [44] Z. Kolská, V. Růžicka, R. Gani, "Estimation of the Enthalpy of Vaporization and the Entropy of Vaporization for Pure Organic Compounds at 298.15 K and at Normal Boiling Temperature by a Group Contribution Method", *Ind. Eng. Chem. Res.* **2005**, *44*, 8436-8454.
- [45] P. S. Beauchamp, R. Marquez, "A General Approach for Calculating Proton Chemical Shifts for Methyl, Methylene and Methine Protons When There Are One or More Substituents within Three Carbons", *J. Chem. Educ.* **1997**, *74*, 1483.
- [46] http://www.ccdc.cam.ac.uk/Lists/ResourceFileList/2014_stats_entries.pdf (Zugriff: 06.01.2014).
- [47] <http://www.rcsb.org/pdb/home/home.do> (Zugriff: 31.03.2015).
- [48] G. Klebe, "Toward a more efficient handling of conformational flexibility in computer-assisted modelling of drug molecules", *Perspect. Drug Discovery Des.* **1995**, *3*, 85-105.
- [49] G. Neudert, G. Klebe, "DSX: A Knowledge-Based Scoring Function for the Assessment of Protein-Ligand Complexes", *J. Chem. Inf. Model.* **2011**, *51*, 2731-2745.
- [50] H. F. G. Velec, H. Gohlke, G. Klebe, "DrugScoreCSDKnowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of Near-Native Ligand Poses and Better Affinity Prediction", *J. Med. Chem.* **2005**, *48*, 6296-6303.
- [51] G. Barnickel, "Molecular Modelling – von der Theorie zur Wirklichkeit", *Chem. Unserer Zeit* **1995**, *29*, 176-185.
- [52] U. Dinur, A. T. Hagler, in *Rev. Comput. Chem.*, John Wiley & Sons, Inc., **2007**, pp. 99-164.
- [53] P. Gerber, K. Müller, "MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry", *J. Comput. Aided Mol. Des.* **1995**, *9*, 251-268.
- [54] C. G. Wermuth, C. R. Ganellin, P. Lindberg, L. A. Mitscher, in *Pure Appl. Chem.*, Vol. **70**, **1998**, p. 1129.
- [55] P. J. Goodford, "A computational procedure for determining energetically favorable binding sites on biologically important macromolecules", *J. Med. Chem.* **1985**, *28*, 849-857.
- [56] M. von Itzstein, W.-Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. V. Phan, M. L. Smythe, H. F. White, S. W. Oliver, P. M. Colman, J. N. Varghese, D. M. Ryan, J. M.


- Woods, R. C. Bethell, V. J. Hotham, J. M. Cameron, C. R. Penn, "Rational design of potent sialidase-based inhibitors of influenza virus replication", *Nature* **1993**, 363, 418-423.
- [57] A. Miranker, M. Karplus, "Functionality maps of binding sites: A multiple copy simultaneous search method", *Proteins* **1991**, 11, 29-34.
- [58] M. Garland R, B. C. David, B. Heinz E, D. Richard A, D. Deborah A, in *Computer-Assisted Drug Design, Vol. 112*, AMERICAN CHEMICAL SOCIETY, **1979**, pp. 205-226.
- [59] P. Labute, C. Williams, M. Feher, E. Sourial, J. M. Schmidt, "Flexible Alignment of Small Molecules", *J. Med. Chem.* **2001**, 44, 1483-1490.
- [60] *Molecular Operating Environment (MOE)*, 2014.09; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2014**.
- [61] M. Nordberg, J. Duffus, D. M. Templeton, in *Pure Appl. Chem.*, Vol. 76, **2004**, p. 1033.
- [62] N. Brown, "Bioisosteres and Scaffold Hopping in Medicinal Chemistry", *Mol. Inform.* **2014**, 33, 458-462.
- [63] G. A. Patani, E. J. LaVoie, "Bioisosterism: A rational approach in drug design", *Chem. Rev. (Washington, D. C.)* **1996**, 96, 3147-3176.
- [64] L. P. Hammett, "Effect of structure upon the reactions of organic compounds. Benzene derivatives", *J. Am. Chem. Soc.* **1937**, 59, 96-103.
- [65] T. Yoshida, M. Shimizu, M. Harada, S. Hitaoka, H. Chuman, "Reassessment of Hammett σ as an effective parameter representing intermolecular interaction energy—links between traditional and modern QSAR approaches", *Bioorg. Med. Chem. Lett.* **2012**, 22, 124-128.
- [66] G. L. Warren, C. W. Andrews, A.-M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D. Wall, J. M. Woolven, C. E. Peishoff, M. S. Head, "A Critical Assessment of Docking Programs and Scoring Functions", *J. Med. Chem.* **2005**, 49, 5912-5931.
- [67] R. D. Taylor, P. J. Jewsbury, J. W. Essex, "A review of protein-small molecule docking methods", *J. Comput. Aided Mol. Des.* **2002**, 16, 151-166.
- [68] H. M. Kumalo, S. Bhakat, M. E. S. Soliman, "Theory and applications of covalent docking in drug discovery: merits and pitfalls", *Molecules* **2015**, 20, 1984-2000, 1917 pp.
- [69] R. Mah, J. R. Thomas, C. M. Shafer, "Drug discovery considerations in the development of covalent inhibitors", *Bioorg. Med. Chem. Lett.* **2014**, 24, 33-39.
- [70] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, "Development and validation of a genetic algorithm for flexible docking", *J. Mol. Biol.* **1997**, 267, 727-748.

- [71] J. Schröder, A. Klinger, F. Oellien, R. J. Marhofer, M. Duszenko, P. M. Selzer, "Docking-based virtual screening of covalently binding ligands: an orthogonal lead discovery approach", *J. Med. Chem.* **2013**, 56, 1478-1490.
- [72] K. Zhu, K. W. Borrelli, J. R. Greenwood, T. Day, R. Abel, R. S. Farid, E. Harder, "Docking Covalent Inhibitors: A Parameter Free Approach To Pose Prediction and Scoring", *J. Chem. Inf. Model.* **2014**, 54, 1932-1940.
- [73] N. London, R. M. Miller, S. Krishnan, K. Uchida, J. J. Irwin, O. Eidam, L. Gibold, P. Cimermančič, R. Bonnet, B. K. Shoichet, J. Taunton, "Covalent docking of large libraries for the discovery of chemical probes", *Nat. Chem. Biol.* **2014**, 10, 1066-1072.
- [74] M. M. Mysinger, B. K. Shoichet, "Rapid Context-Dependent Ligand Desolvation in Molecular Docking", *J. Chem. Inf. Model.* **2010**, 50, 1561-1573.
- [75] X. Ouyang, S. Zhou, C. T. Su, Z. Ge, R. Li, C. K. Kwoh, "CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints", *J. Comput. Chem.* **2013**, 34, 326-336.
- [76] P. M. Morse, "Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels", *Phys. Rev.* **1929**, 34, 57-64.
- [77] W. Gerschlager, G. Baumgart, *Alzheimer: die Krankheit des Vergessens*, Maudrich, **2007**.
- [78] R. Bischoff, M. Simm, R. A. Zell, *Demenzforschung im Fokus*, Bundesministerium für Bildung und Forschung, **2004**.
- [79] I.-C. Stancu, B. Vasconcelos, D. Terwel, I. Dewachter, "Models of β -amyloid induced Tau-pathology: the long and "folded" road to understand the mechanism", *Mol. Neurodegener.* **2014**, 9, 51/51-51/28.
- [80] H. Braak, E. Braak, D. Yilmazer, R. A. de Vos, E. N. Jansen, J. Bohl, "Pattern of brain destruction in Parkinson's and Alzheimer's diseases", *J. Neural Transm.* **1996**, 103, 455-490.
- [81] L. Gravit, "Drugs: A tangled web of targets", *Nature* **2011**, 475, S9-S11.
- [82] D. R. Thal, U. Rub, C. Schultz, I. Sassin, E. Ghebremedhin, K. Del Tredici, E. Braak, H. Braak, "Sequence of A β -protein deposition in the human medial temporal lobe", *J. Neuropathol. Exp. Neurol.* **2000**, 59, 733-748.
- [83] F. M. LaFerla, K. N. Green, S. Oddo, "Intracellular amyloid-[β] in Alzheimer's disease", *Nat. Rev. Neurosci.* **2007**, 8, 499-509.
- [84] R. Medeiros, D. Baglietto-Vargas, F. M. LaFerla, "The Role of Tau in Alzheimer's Disease and Related Disorders", *CNS Neurosci. Ther.* **2011**, 17, 514-524.
- [85] M. P. Mazanetz, P. M. Fischer, "Untangling tau hyperphosphorylation in drug design for neurodegenerative diseases", *Nat. Rev. Drug Discov.* **2007**, 6, 464-479.

- [86] S.-R. Ryoo, H. K. Jeong, C. Radnaabazar, J.-J. Yoo, H.-J. Cho, H.-W. Lee, I.-S. Kim, Y.-H. Cheon, Y. S. Ahn, S.-H. Chung, W.-J. Song, "DYRK1A-mediated Hyperphosphorylation of Tau: A Functional Link Between Down Syndrome and Alzheimer Disease", *J. Biol. Chem.* **2007**, *282*, 34850-34857.
- [87] E. Kalbe, P. Calabrese, S. Schwalen, J. Kessler, "The Rapid Dementia Screening Test (RDST): a new economical tool for detecting possible patients with dementia", *Dement. Geriatr. Cogn. Disord.* **2003**, *16*, 193-199.
- [88] W. E. Klunk, H. Engler, A. Nordberg, Y. Wang, G. Blomqvist, D. P. Holt, M. Bergström, I. Savitcheva, G.-F. Huang, S. Estrada, B. Ausén, M. L. Debnath, J. Barletta, J. C. Price, J. Sandell, B. J. Lopresti, A. Wall, P. Koivisto, G. Antoni, C. A. Mathis, B. Långström, "Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B", *Ann. Neurol.* **2004**, *55*, 306-319.
- [89] W. Zhen, H. Han, M. Anguiano, C. A. Lemere, C.-G. Cho, P. T. Lansbury, "Synthesis and Amyloid Binding Properties of Rhenium Complexes: Preliminary Progress Toward a Reagent for SPECT Imaging of Alzheimer's Disease Brain", *J. Med. Chem.* **1999**, *42*, 2805-2815.
- [90] M. Koronyo-Hamaoui, Y. Koronyo, A. V. Ljubimov, C. A. Miller, M. K. Ko, K. L. Black, M. Schwartz, D. L. Farkas, "Identification of amyloid plaques in retinas from Alzheimer's patients and noninvasive in vivo optical imaging of retinal plaques in a mouse model", *NeuroImage* **2011**, *54*, Supplement 1, S204-S217.
- [91] L. Gasparini, R. Anthony Crowther, K. R. Martin, N. Berg, M. Coleman, M. Goedert, M. G. Spillantini, "Tau inclusions in retinal ganglion cells of human P301S tau transgenic mice: Effects on axonal viability", *Neurobiol. Aging* **2011**, *32*, 419-433.
- [92] V. Labonté, "Augen auf - Alzheimer Früherkennung", *Laborjournal* **2010**, *4/2010*, 14.
- [93] G. G. Stokes, "On the Change of Refrangibility of Light", *Phil. Trans. R. Soc. Lond.* **1852**, *142*, 463-562.
- [94] S. Burgold, T. Bittner, M. Dorostkar, D. Kieser, M. Fuhrmann, G. Mitteregger, H. Kretzschmar, B. Schmidt, J. Herms, "In vivo multiphoton imaging reveals gradual growth of newborn amyloid plaques over weeks", *Acta Neuropathol.* **2011**, *121*, 327-335.
- [95] fluorophores.org (Zugriff: 08.04.2015).
- [96] P. W. Atkins, J. d. Paula, *Physikalische Chemie*, 4 ed., Wiley-VCH, **2006**.
- [97] M. F. Cordeiro, L. Guo, K. M. Coxon, J. Duggan, S. Nizari, E. M. Normando, S. L. Sensi, A. M. Sillito, F. W. Fitzke, T. E. Salt, S. E. Moss, "Imaging multiple phases of neurodegeneration: a novel approach to assessing cell death in vivo", *Cell Death Dis.* **2010**, *1*, 1e-11e.

- [98] W. E. Klunk, B. J. Bacsikai, C. A. Mathis, S. T. Kajdasz, M. E. McLellan, M. P. Frosch, M. L. Debnath, D. P. Holt, Y. Wang, B. T. Hyman, "Imaging Abeta plaques in living transgenic mice with multiphoton microscopy and methoxy-X04, a systemically administered Congo red derivative", *J. Neuropathol. Exp. Neurol.* **2002**, *61*, 797-805.
- [99] W. Becker, U. Soppa, F. J. Tejedor, "DYRK1A: A Potential Drug Target for Multiple Down Syndrome Neuropathologies", *CNS Neurol. Disord. Drug Targets* **2014**, *13*, 26-33.
- [100] W. Becker, W. Sippl, "Activation, regulation, and inhibition of DYRK1A", *FEBS J.* **2011**, *278*, 246-256.
- [101] E. J. Kim, J. Y. Sung, H. J. Lee, H. Rhim, M. Hasegawa, T. Iwatsubo, D. S. Min, J. Kim, S. R. Paik, K. C. Chung, "Dyrk1A Phosphorylates α -Synuclein and Enhances Intracellular Inclusion Formation", *J. Biol. Chem.* **2006**, *281*, 33250-33257.
- [102] J. E. Kang, S. A. Choi, J. B. Park, K. C. Chung, "Regulation of the proapoptotic activity of huntingtin interacting protein 1 by Dyrk1 and caspase-3 in hippocampal neuroprogenitor cells", *J. Neurosci. Res.* **2005**, *81*, 62-72.
- [103] I. Ferrer, M. Barrachina, B. Puig, M. Martínez de Lagrán, E. Martí, J. Avila, M. Dierssen, "Constitutive Dyrk1A is abnormally expressed in Alzheimer disease, Down syndrome, Pick disease, and related transgenic models", *Neurobiol. Dis.* **2005**, *20*, 392-400.
- [104] R. Kimura, K. Kamino, M. Yamamoto, A. Nuripa, T. Kida, H. Kazui, R. Hashimoto, T. Tanaka, T. Kudo, H. Yamagata, Y. Tabara, T. Miki, H. Akatsu, K. Kosaka, E. Funakoshi, K. Nishitomi, G. Sakaguchi, A. Kato, H. Hattori, T. Uema, M. Takeda, "The DYRK1A gene, encoded in chromosome 21 Down syndrome critical region, bridges between β -amyloid production and tau phosphorylation in Alzheimer disease", *Hum. Mol. Gen.* **2007**, *16*, 15-23.
- [105] Y. L. Woods, P. Cohen, W. Becker, R. Jakes, M. Goedert, X. Wang, C. G. Proud, "The kinase DYRK phosphorylates protein-synthesis initiation factor eIF2B ϵ at Ser539 and the microtubule-associated protein tau at Thr212: potential role for DYRK as a glycogen synthase kinase 3-priming kinase", *Biochem. J.* **2001**, *355*, 609-615.
- [106] D. Frost, B. Meechoovet, T. Wang, S. Gately, M. Giorgetti, I. Shcherbakova, T. Dunckley, " β -Carboline Compounds, Including Harmine, Inhibit DYRK1A and Tau Phosphorylation at Multiple Alzheimer's Disease-Related Sites", *PLoS One* **2011**, *6*, e19264.
- [107] J. Shi, T. Zhang, C. Zhou, M. O. Chohan, X. Gu, J. Wegiel, J. Zhou, Y.-W. Hwang, K. Iqbal, I. Grundke-Iqbal, C.-X. Gong, F. Liu, "Increased Dosage of Dyrk1A Alters

- Alternative Splicing Factor (ASF)-regulated Alternative Splicing of Tau in Down Syndrome", *J. Biol. Chem.* **2008**, 283, 28660-28669.
- [108] S.-R. Ryoo, H.-J. Cho, H.-W. Lee, H. K. Jeong, C. Radnaabazar, Y.-S. Kim, M.-J. Kim, M.-Y. Son, H. Seo, S.-H. Chung, W.-J. Song, "Dual-specificity tyrosine(Y)-phosphorylation regulated kinase 1A-mediated phosphorylation of amyloid precursor protein: evidence for a functional link between Down syndrome and Alzheimer's disease", *J. Neurochem.* **2008**, 104, 1333-1344.
- [109] Y. S. Ryu, S. Y. Park, M.-S. Jung, S.-H. Yoon, M.-Y. Kwen, S.-Y. Lee, S.-H. Choi, C. Radnaabazar, M.-K. Kim, H. Kim, K. Kim, W.-J. Song, S.-H. Chung, "Dyrk1A-mediated phosphorylation of Presenilin 1: a functional link between Down syndrome and Alzheimer's disease", *J. Neurochem.* **2010**, 115, 574-584.
- [110] Y. Ogawa, Y. Nonaka, T. Goto, E. Ohnishi, T. Hiramatsu, I. Kii, M. Yoshida, T. Ikura, H. Onogi, H. Shibuya, T. Hosoya, N. Ito, M. Hagiwara, "Development of a novel selective inhibitor of the Down syndrome-related kinase Dyrk1A", *Nat. Commun.* **2010**, 1, 86.
- [111] B. Smith, F. Medda, V. Gokhale, T. Dunkley, C. Hulme, "Recent Advances in the Design, Synthesis, and Biological Evaluation of Selective DYRK1A Inhibitors: A New Avenue for a Disease Modifying Treatment of Alzheimer's?", *ACS Chem. Neurosci.* **2012**, 3, 857-872.
- [112] N. Göckler, G. Jofre, C. Papadopoulos, U. Soppa, F. J. Tejedor, W. Becker, "Harmine specifically inhibits protein kinase DYRK1A and interferes with neurite formation", *FEBS Journal* **2009**, 276, 6324-6337.
- [113] J. Bain, L. Plater, M. Elliott, N. Shpiro, C. J. Hastie, H. McLauchlan, I. Klevernic, J. S. C. Arthur, D. R. Alessi, P. Cohen, "The selectivity of protein kinase inhibitors: a further update", *Biochem. J.* **2007**, 408, 297-315.
- [114] T. Herraiz, D. González, C. Ancín-Azpilicueta, V. J. Arán, H. Guillén, "β-Carboline alkaloids in Peganum harmala and inhibition of human monoamine oxidase (MAO)", *Food Chem. Toxicol.* **2010**, 48, 839-845.
- [115] <http://www.chemcomp.com/journal/svl.htm> (Zugriff: 02.04.2015).
- [116] http://daylight.com/dayhtml_tutorials/languages/smirks/index.html (Zugriff: 09.04.2015).




Christoph Scholz
Kaiserstraße 108
61169 Friedberg

29.04.2015

Erklärung

Ich erkläre hiermit, dass ich meine Dissertation selbstständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

(Christoph Scholz)



Christoph Scholz
Kaiserstraße 108
61169 Friedberg

29.04.2015

Erklärung

Ich erkläre hiermit, noch keinen Promotionsversuch unternommen zu haben.

(Christoph Scholz)

Lebenslauf



Name:	Christoph Scholz
Geburtsdatum:	07. März 1985
Geburtsort:	Potsdam
Staatsangehörigkeit:	deutsch
Familienstand:	ledig, keine Kinder

BERUFLICHE TÄTIGKEITEN

10/2012 - 06/2015	Wissenschaftlicher Mitarbeiter der TU Darmstadt, Institut für Organische Chemie und Biochemie, Prof. Dr. B. Schmidt
--------------------------	---

HOCHSCHULSTUDIUM

- 10/2012 - heute** Promotion (Dr. rer. nat.) an der TU Darmstadt, Institut für Organische Chemie und Biochemie, Prof. Dr. B. Schmidt
- 04/2010 - 08/2012** Chemiestudium (Master of Science Chemie) an der TU Darmstadt, Schwerpunktfach: Organische Chemie, Hauptfächer: Biochemie und Makromolekulare Chemie
- Thema der Thesis: „*Ligandenbasierte Untersuchung eines induced-fit-Mechanismus und Synthese von Inhibitoren des 20S-Proteasoms.*“
Betreuer der Thesis: Prof. Dr. Boris Schmidt, Institut für Organische Chemie und Biochemie
Abschluss: Master of Science (sehr gut)
- 10/2006 - 03/2010** Chemiestudium (Bachelor of Science Chemie) an der TU Darmstadt
- Thema der Thesis: „*Amidinfunktionalisierte Stilben- und Oxadiazolderivate für bildgebende und biozide Applikationen.*“
Betreuer der Thesis: Prof. Dr. Boris Schmidt, Institut für Organische Chemie und Biochemie
Abschluss: Bachelor of Science (gut)

STUDIENBEGLEITENDE TÄTIGKEITEN

- 2006 - 2012** Chemielaborant bei der Fresenius Medical Care AG, Steinmühlstraße 24, 61352 Bad Homburg

BERUFSAUSBILDUNG

- 09/2003 - 07/2006** Ausbildung zum Chemielaboranten bei der Fresenius AG, Else-Kröner-Straße 1, 61352 Bad Homburg

SCHULAUSBILDUNG

- 02/2004 – 11/2005** Ausbildungsbegleitende Fachoberschule, Paul-Ehrlich-Schule Frankfurt/M.
Abschluss: Fachhochschulreife
- 02/2001 - 07/2002** Staatliches Eifel-Gymnasium, Neuerburg
Abschluss: Mittlere Reife
- 08/1995 - 01/2001** Max-Planck-Gymnasium, Rüsselsheim

ZIVILDIENTST

- 12/2002 – 09/2003** DRK-Blutspendedienst, Frankfurt/M.